Infant learning rate within conditioning procedures as a predictor of subsequent psychological development

A.J.M. Goossens

INFANT LEARNING RATE WITHIN CONDITIONING PROCEDURES AS A PREDICTOR OF SUBSEQUENT PSYCHOLOGICAL DEVELOPMENT

INFANT LEARNING RATE WITHIN CONDITIONING PROCEDURES AS A PREDICTOR OF SUBSEQUENT PSYCHOLOGICAL DEVELOPMENT

Een wetenschappelijke proeve op het gebied van de Sociale Wetenschappen

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit te Nijmegen,
volgens besluit van het college van decanen
in het openbaar te verdedigen op
dinsdag 5 juni 1990
des namiddags te 1.30 uur precies

door

Aloysius Josephus Maria Goossens

geboren op 20 maart 1951 te Gemert

Promotor: Prof. dr. J.J. Dumont

Co-promotores: Dr. G.E. Lancioni (Ph.D) (R.U. Leiden)
Dr. J.H.L. Oud

Lay-out en zetwerk: Infobever, Nijmegen
ISBN 90-9003409-9

aan Angret
aan mijn ouders

Acknowledgements

The research reported in this thesis was supported by a Grant from the Onderzoekspool of the Katholieke Universiteit Nijmegen, Faculty of the Social Sciences, Grant SW 1/81.

My thanks go especially to the subjects and their parents, with their ready willingness to participate, which made the study possible and its execution a pleasure. In addition, I wish to express my appreciation to Drs Boon, Hoevenaars, van der Zee, van Wieringen and their medical staff from the divisions of Neonatlogy and Intensive Care of the Sint Radboud University Hospital and Canisius-Wilhelmina Hospital, both in Nijmegen, for subject selection and recruitment.

I gratefully acknowledge Dr. G.E. Lancioni in preparation of the research proposal, supervision during the execution of the study and preparation of the manuscript. I also wish to thank Prof. Dr. J.J. Dumont and Dr. J.H.L. Oud for their obliging support during the preparation of the manuscript. The assistance of Drs. W. de Bruijn in the data collection was greatly appreciated.

In addition, I wish to thank M. Hermans and A. Findhammer for typing parts of the manuscript, M. Schakenraad for his contribution to the English text, and R. Timmer, research-coordinator, for his help in the execution of the study. Moreover, my thanks are due to H. Creemers who illustrated the front page and to the many persons who contributed in any stage of this work and are not mentioned.

I also wish to thank the collaborators of the technical departments of the Faculty of Social Sciences (electronics and mechanics) for their contribution in designing and constructing of all the equipment.

The author is now at the Department SIBD, Instituut voor Doven, Theerestraat 42, 5271 GD Sint Michielsgestel, Holland.

Contents

Introduction

Assessment of cognitive skills in very young children is considered important with respect to the estimation of present as well as future capacities. Especially relevant seems the aforementioned goal of assessment in high-risk and handicapped children. Generally one prefers to assess as early as possible. Early assessment is mostly aimed at the measurement of present levels of functioning upon which eventual intervention is based and at evaluation of the effects of such interventions, directed at preventing or overcoming developmental problems. In addition, such assessment could also contribute significantly to the design of (individual) intervention plans.

From the beginning of the century many efforts have been undertaken to pursue the aforementioned goals. All these efforts have resulted in standardized infant tests, analogous to those developed for children and adults, and more standardization in neurological and neurophysiological assessments. Although these tests and assessments have led to more accurate estimates of infant status, prediction of future developmental level showed to be rather disappointing. In addition, there was a heightening interest and need in reliable assessment of cognitive functions and processes playing a role in the developing organism.

The aforementioned considerations and others have evolved in a wide field of research on infant cognitive functioning (e.g., sensory processing and learning) and infant assessment. The knowledge within this area increased vastly and, perhaps more importantly, the view of the infant as having a passive role in the first phase of development was increasingly abandoned. Now the infant is seen as an active participator interacting in a complicated environment.

The historical background of the aforementioned developments and some relevant literature describing them unlock the present report. That is, a brief history and background of infant testing appears in Chapter 1 of this paper. Several procedures, old and new, are outlined. In Chapter 2 recent developments in the area of traditional assessment procedures, including neurological measurements, are described. In addition, several important test instruments are delineated in some more detail, for example, the Bayley test and Prechtl's neurological examination.

Chapter 3 provides an overview of new assessment strategies, possible reasons for the rise in interest in them, and problems they had (and still have) to overcome. More specifically, neurophysiological, environmental, attention, some specific measurements, and conditioned learning assessments are sketched in sections 3.2, 3.3, 3.4, and 3.5, respectively. In addition, some results with regard to the validity of the aforementioned procedures, which pertain to the assessment of developmental potential or prediction of future development, are discussed in these sections.

In Chapter 4 an experiment is described, aimed at the assessment of learning rate in infants. The rationale, method and procedures of this experiment, in which classical and operant conditioning procedures as well as home observations were applied at the ages of 4 and 6 months, are outlined in sections 4.1 and 4.2. The learning rates, derived from these conditioning procedures, and the observation measures are used to predict subsequent development, measured by an operant conditioning procedure and the Bayley Scales of Infant Development at the age of 18 months. Furthermore, in section 4.3 the results of the aforementioned experiment are presented, that is, the data are described for the two subgroups of subjects applied (a group of high-risk and handicapped subjects and a group of normal subjects) and comparisons are made between them. In addition, the data are analysed by means of simple correlational and multiple multivariate statistical procedures. Moreover, the groups of successful and unsuccessful high-risk subjects (successfulness is defined in relation to the conditioning procedures) are compared on their performances on the Bayley Mental and Motor Scales. Finally, all subjects are classified according to performance on the Bayley Mental Scale, that is, through discriminant analysis group membership is predicted using a selected set of predictive measures.

In conclusion, in section 4.4 the summary and conclusions as well as some implications of the present study are provided. In addition, several suggestions for future research are made.

1. History and background of infant assessment

1.1. General outline

Around the turn of the century, interest in individual differences in higher mental functions and in strategies to measure them started to be developed. This led to the introduction of the first intelligence test for children, that is, the scale of Binet and Simon in France in 1905. The test was devised to detect those children who were too dull to profit from ordinary schooling. It contained items ranked in order of level of difficulty and was accompanied by relatively precise instructions for administration. In 1916, Terman standardized the same scale in the USA: the Stanford-Binet. The latter test was later revised repeatedly. This early period was followed by a time of relative indifference, which protracted itself until the 1930s. At that time, the issue of assessment became very popular and a variety of infant tests was presented, most of them in the USA (cf. Lewis, 1976; Sattler, 1988).

The most important tests were the Gesell Developmental Schedules, devised under the direction of A. Gesell (1940-1947), the Viennese Test Series presented by C. Buhler, H. Hetzer and their colleagues in 1932 in Austria, Shirley's Study of the First Two Years of Infant's Life by M. Shirley in 1933, the California First Year Mental Scale by N. Bayley in 1933, and the Iowa Tests for Young Children by E. Fillmore in 1936. The assessment instruments were primarily aimed at describing behavioral development in longitudinal infant studies. In addition, they were developed and used for diagnosing abnormal development and predicting subsequent development. The aforementioned instruments will be described in more detail in the following section.

1.2. Assessment instruments

The Gesell Development Schedules described behavioral development in children from 1 month through 6 years. The schedules provided standardized procedures for observation and evaluation of such development. The items, obtained after systematic and careful observations and examinations of hundreds of children, were grouped into four subtests, covering the areas of motor, adaptive, language and personal-social development.

The Viennese Test Series covered, in the 1932 version, the age range from birth to 5 years. Items were designed to measure sensory development, bodily movements, social behavior (including language), learning (and imitation), manipulation of objects, and thinking processes.

Shirley's Study of the First Two Years of Infant's Life offered a comprehensive description of behavioral development during the first years of life. Shirley applied weekly exams which concentrated on physical development and reactions to sensory stimuli including verbal instructions and pictures. The test was never used widely.

An influential scale turned out to be the California First Year Mental Scale (by Bayley in the Berkeley Growth Study, in 1933) used for children from birth through 3 years. The scale was composed of 185 items, original as well as from other tests. Bayley attempted to find items that represented mental functioning and that would predict later intelligence. The items (showing age-related composition) tapped motor maturation, eye-hand coordination, adaptive behavior, response to sound, visual maturation, language comprehension and production, and social responsiveness.

The Iowa Tests for Young Children could be applied to children from birth to 3 years of age. The tests contained items similar to those of the aforementioned tests. Fillmore attempted to include only those items that predicted later IQ.

By the middle of the century the aforementioned tests were used mainly for diagnostic purposes, that is, differentiation of normal from abnormal mental development and assessment of mental capabilities. The tests were used for evaluation of children for adoption, for admission to special schools or programs, as well as for physically handicapped children. At the same time many studies sought to investigate the tests with regard to standardization, internal consistency and test-retest reliability. In addition, the predictive validity from infancy to early childhood was carefully scrutinized. The results of these investigations were satisfactory as far as the internal consistency of the tests was concerned. Test-retest reliability, however, was fairly low, while the predictive validity showed even more disappointing results (McCall, Hogarty, & Hurlburt, 1972; McCall, 1982). Within this context, some authors continued on the use of such instruments as diagnostic tools. Others, however, attempted to modify the existing tools, notably Shotwell and Gilliland (1943) and Cattell (1966) in the USA, and Griffiths in Great Britain (1954). In developing their assessment instruments, they relied heavily on available items and strategies. Their main purpose was to devise tests with better psychometric properties and to extend measurement towards early infancy and later childhood.

Shotwell and Gilliland developed a test, based on items from the Gesell Scales, for infants from 4 to 12 (Test A) and 13 to 36 (Test B) weeks of age. This test was never researched nor used extensively.

Cattell devised a test (a downward revision of the Stanford-Binet, covering the ages from 3 to 30 months) that used several rather strict criteria with regard to item-content and responses required. For example, one criterion was that the items showed a regular increase in percentage of success, across subjects of increasing age. Administration and scoring were precisely defined and data were collected for norms.

Griffith's Mental Development Scale consisted of five subscales, that is, locomotor, personal-social, an extensive subscale on hearing and speech, hand and eye, and performance. It was intended for assessment of innate cognitive ability in infants from 2 weeks to 2 years.

1.3. Summary and discussion

The overall characteristics of all the tests (old and new) were the broad range of behaviors they represented, the reliance on spontaneous responding to the presentation of environmental cues and elicited reflexes or motor behaviors (Gaussen & Stratton, 1985). The broadness of the tests reflected the idea that a general intelligence factor was the basis of mental development (Fagan, 1984). That is, it was assumed that this factor provided developmental continuity and would manifest itself in the performance on the aforementioned broad range of behaviors tested. The set of basic processes underlying the general factor and the resulting knowledge, however, was not identified or measured directly.

In the early sixties most researchers seemed to realize that long-term prediction of general intellectual ability, at least solely via traditional assessment procedures, was unrealistic. Despite such realization, a revival of interest in infants and infant development was set in motion. Two research streams (not always separable or independent) became predominant. The first was a new wave of interest in improving traditional assessment procedures, which involved renewed attention for existing as well as the development of new tests (infant tests). In the same stream another approach can be distinguished, which in essence also comprises traditional assessment procedures, that is, neurological measurements. This latter field of assessment is discussed in a separate subsection of the chapter, because it is supposed to be a completely different aspect of assessment. That is, neurological measurements apply different research methodologies which are directed at the study of central nervous system functions. The second stream comprised what could be called the development of new assessment strategies. These two streams will be discussed in the next two chapters of the paper (Chapters 2 and 3).

2. Refinement and extension of traditional assessment procedures: Infant tests and neurological measurements

2.1. Infant tests

2.1.1. Outline

Refinements of traditional instruments were considered relevant because of the acknowledged need for reliable and comprehensive instruments to determine the developmental level of both normal and abnormal infants. Extension of the age range was also deemed necessary: downwards for assessing preterm and term neonates, upwards for covering the age range onto school age (Francis-Williams, 1977; Thomas, 1970; Yang, 1979).

The following influential instruments, which will be mentioned in the next section, were presented in the period from the early sixties. The Bayley Scales of Infant Development (BSID), the Graham/Rosenblith Behavioral Examination of Neonates, the Neonatal Behavioral Assessment Scale (NBAS), The Assessment of Preterm Infant's Behavior (APIB), the Griffiths Scale, the Gesell Developmental Scales, the Denver Developmental Screening Test (DDST), and the Infant Psychological Development Scales (IPDS). Some earlier versions of these tests were described in the previous section of this paper.

2.2.2. Assessment instruments

The Bayley Scales of Infant Development is one of the most widely used instruments, not only in the USA, but also abroad (Bayley, 1967, 1969). The Bayley is essentially a revision of the so-called California Scales, and is aimed at providing measures of the developmental level of children from 1 month to 2 1/2 years. A Dutch version of the Bayley, with norms based on a national sample, is available (BOS 2-30, Meulen & Smrkovsky, 1983). The Bayley covers mental and motor areas. The mental area includes items on adaptability or learning, sensory acuity, and fine motor coordination (e.g., orientation to a sound, piling cubes). The motor area is composed of items dealing with body coordination (e.g., independent sitting, standing on one foot). In addition, scoring of the Infant Behavior Record (IBR), based on the behavior presented during testing, can be part of an exam. A complete exam, not including the IBR, takes from 1/2 to 1 hour.

Another popular instrument is the Graham/Rosenblith Behavioral Examination of Neonates (Rosenblith, 1961a, 1961b, 1975, 1979). The test was originally designed by Graham, Matarazzo, and Caldwell (1956) to differentiate normal newborns from those with risk of brain injury. This continued to be the main goal of the scale. The test covers the following areas: motor and tactile-adaptive (requiring defensive responses). In addition, some items are used which have to be rated (e.g., irritability, light sensitivity). The test can be administered in 20-30 minutes.

The Neonatal Behavioral Assessment Scale (NBAS) (Brazelton, 1973; Lancioni, Horowitz, & Sullivan, 1980) was devised for normal newborns. This scale includes reflex items for assessment of neurological intactness. Other, more global behavioral dimensions, are rated, for example, attractiveness (e.g., availability for social interaction) and need for stimulation (e.g., toleration of handling and stimulation). Finally and most importantly, 26 items assess the newborn's interactive behavioral repertoire. The examination takes about 20 to 30 minutes, not including scoring. Als and her colleagues (Als & Duffy, 1983; Als, Lester, Tronick, & Brazelton, 1982; Duffy & Als, 1983) developed the Assessment of Preterm Infants' Behavior (APIB). They relied heavily on the NBAS. The APIB is especially designed for preterm infants.

The Griffiths Scale, also extensively applied, particularly in Europe, is developed to distinguish between normal and handicapped children. Its age range, starting at 1 month and originally ending at 24 months, was extended to 8 years of age (Griffiths, 1970). The test, the first version of which was briefly described above, takes little note of the quality of the infant's response, his alertness and his interest in the surroundings.

The Gesell Developmental Scales (Gesell & Amatruda, 1947) can be applied for children of 1 month through 6 years. Since its first presentation (1947) the structure of the Scale was not significantly altered (see above). However, the age norms were changed and based on a much more extended sample of subjects (Knobloch & Pasamanick, 1974; Knobloch, Stevens, & Malone, 1980).

The Denver Developmental Screening Test (Frankenburg & Coons, 1983; Frankenburg & Dodds, 1967; Cools & Hermanns, 1977) is frequently applied. It is used in the first six years of life (from 2 weeks to 6 years) and aimed specifically at early identification of serious developmental delays. It gives scores in four different areas of behavior, that is, gross motor, fine motor-adaptive, language, and personal-social individual behavior.

The Infant Psychological Development Scales, developed by Uzgiris and Hunt (1975), is an assessment procedure based on Piagetian theory. It is suited for children of 1 to 24 months of age (see also Uzgiris, 1976). The assessment comprises six scales, that is, visual pursuit and permanence of objects, development of means for obtaining desired environmental events, imitation, operational causality, construction of object relations in space, and finally, the development of schemes for relating to objects.

2.1.3. Summary and discussion

Several of the aforementioned assessment procedures show similarities among them (e.g., the Bayley and the Griffiths), mainly with regard to the choice of items and thus the responses required. However, also differences exist, which are generally shown in the area of scoring and interpretation (cf. the Uzgiris/Hunt Scales). The standards and norms may also differ across the different scales.

Despite the fact that the scales are refined or extended, the long-term prediction of intelligence, from a very young age, is still poor or yet to be determined. Correlations of .40 to negligible were found when infant tests were applied and test performances measured at 6 months and 36 months of age were related (Kopp & McCall, 1982; McCall, 1982, 1983; Yang, 1979). Only for infants functioning at the extreme low end of the developmental continuum, prediction of outcome seems to be reliable (Honzik, 1976).

From various scientific backgrounds different reasons have been suggested for the poor predictive validity of the aforementioned assessment instruments. The first and most widely stressed reason is the differential influence of environmental variables on development. Variables such as physical and social environment and cognitive stimulation or motivation are mentioned (cf. Wachs & Gruen, 1982, for a discussion of this topic). Other common reasons are normally prevailing discontinuities in developmental trends, unanticipated degenerative diseases (Brooks & Weinraub, 1976), and the assumed developmental plasticity of the central nervous system (Gollin, 1981; Lipton, 1976). The highly unpredictable influences of the latter factors may play a big role in abnormal or high-risk populations for which standardized assessment instruments hardly exist. Generally, intelligence tests are constructed on the assumption that a general intelligence factor exists. This factor is assumed to be genetically fixed and relatively unaffected by environmental influences, providing the basis for cognitive and behavioral continuity. Such an assumption may be too simple and may not hold for a rapidly developing organism (Fagan, 1984; Fagan & Singer, 1983; Kopp & McCall, 1982; Rose, 1981; Sattler, 1988; Ulvund, 1984). Even holding the assumption true, it could be that the tests measure behaviors that are not significant representations of infant status and developmental potential. That is, simple sensorimotor skills may bear little relation to concurrent or subsequent cognitive functioning. Thus infant tests cover a period in which important qualitative changes occur in the child's development (Caron, Caron & Glass, 1983; Fagen, Ohr, Fleckenstein, & Singer, 1987; Gaussen & Stratton, 1985; McCall, 1983; Yarrow, Klein, Lomonaco, & Morgan, 1975). In addition to these theoretical points of discussion, in many studies predictor and outcome measures are obtained by different instruments. Even when the same test is applied at different ages the resulting scores are based on different test-items, which assess from sensorimotor to more cognitively oriented functions. These and other measurement factors give rise to serious methodological questions (Sattler, 1988).

2.2. Neurological measurements

2.2.1. Outline

Neurological measurements are traditional assessment procedures for newborns. They are mostly applied by child-neurologists and pediatricians in hospital settings. The measurements are primarily designed to assess the level of maturity or the condition of the central nervous system (Dubowitz, Dubowitz, & Goldberg, 1970; Lubchenko, 1983; Parmelee & Michaelis, 1971; Prechtl & Beintema, 1977). This condition is expected to provide information about the level of functioning in motor, sensory, cognitive and behavioral areas. In case of adverse conditions, temporary or lasting, not only present but also future functioning could be affected (Prechtl & Beintema, 1977; Prechtl, 1982, 1983). The assessment may also guide decisions concerning treatment or special care.

Neurological investigations may take place in the neonatal period, but follow-ups are generally planned to provide information on the persistence of abnormal or doubtful responses (Prechtl & Beintema, 1977; Prechtl 1983; Saint-Anne Dargassies, 1977, 1979). Neurological assessment may also be applied in addition to other assessments (e.g., psychological) in order to investigate relationships with them and thus improve prediction of later development (Parmelee & Michaelis, 1971). The neurological measurements, particularly the screening tests, mainly consist of observation of simple reflexes and do not require the active cooperation of the infant (St. Clair, 1978).

Some of the procedures which are most commonly applied and are highly systematic and detailed are mentioned in the following section. They are the Dubowitz, the Prechtl and the Parmelee schemes.

2.2.2. Assessment instruments

The Dubowitz scheme (Dubowitz et al., 1970; Self & Horowitz, 1979) is very widely applied to estimate gestational age of a newborn, especially in case of suspected prematurity or low birthweight. The assessment intends to differentiate small-for-date, pre-term, and full-term infants. The latter characteristics are generally incorporated in risk-indexes and used to indicate different levels of abnormality (Lubchenko, 1983; Self & Horowitz, 1979). The scoring system is based on neurological and physical characteristics of the neonate, such as posture and skin texture. Its value for prediction of later cognitive development, however, does not seem to be very meaningful. For example, in a study applying this instrument, among several other perinatal and infant measures, extremely weak predictions of subsequent IQ and language measures were found (Bee et al., 1982).

The Prechtl and Parmelee schemes have been applied and investigated very extensively. The Prechtl optimality method (Prechtl & Beintema, 1977) includes items on general information (e.g., method of feeding, drugs administered), an observation period for assessment of state, and an extensive examination of physical characteristics such as posture or skin color and spontaneous and elicited reflexes. Instead of relying on disease entities and pathological signs which may or may not have occurred in the history of individuals, optimality is defined according to a predefined set of strict criteria. Optimal refers to the best possible condition carrying the least risk of mortality and morbidity. Application of the scheme results in a score which indicates the degree of optimality: optimality versus reduced optimality (Prechtl, 1983). Several reports point to the usefulness of the method for diagnostic and prognostic purposes.

For instance, Bierman-van Eendenburg, Jurgens-van der Zee, Olinga, Huisjes, and Touwen (1981) found that the instrument, which they applied on 1507 newborns in the Groningen Perinatal Project, was sensitive in detecting neurological dysfunction. In fact, they re-examined, at 18 months of age, 80 infants who had been neurologically abnormal in the neonatal period and 80 (normal) controls. Of the 80 abnormal infants, one had died at the age of 3 months and 13 other remained abnormal at 18 months. However, only two of the control infants had mild abnormalities at 18 months.

Hadders-Algra, Huisjes, and Touwen (1988a, 1988b) assessed a sample of an extension of the cohort in the Groningen Perinatal Project (comprising 3162 subjects) at the age of 9 years. The sample consisted of all neurologically abnormal infants (n = 160), a group of neonatally mildly abnormal infants (n=322), and a random sample of normal neonates (n=322). They found that none of the subjects classified normal at the neonatal examination, according to the technique of Prechtl, were neurologically abnormal at follow-up (at 9 years of age). Of the neurologically abnormal neonates 10% were neurologically handicapped at follow-up. They found early postnatal factors, low social class, and the presence of interval complications to be related with deviant outcome. In addition, a high rate of children who recovered from their deviant neurological condition at birth was reported. The same authors found in a second study (Hadders-Algra et al., 1988b) that neonatal neurological abnormalities contributed to cognitive (reading, spelling, and arithmetic) and behavioral difficulties at 9 years of age.

Njiokiktjien and Kurver (1980) divided a group of normal newborns in neurologically "optimal" and "suboptimal" groups, according to the Prechtl exam. Differences between the groups (unfavorable for the suboptimal) were found on behavioral scores like emotional lability, particular patterns of sleep and short attention span. These scores were derived from observations in the laboratory and from questionnaires filled in by the parents when their children were 12, 21 and 36 months old.

Parmelee's exam is applied in many studies, generally in concomitance with other assessment procedures. The neurological examination is short (10 minutes) and easily administered. It assesses organized patterns of behavior, myotatic tonus, and states of arousal in infants at 40 weeks gestational age (Parmelee, Kopp, & Sigman, 1976).

As an example of a well designed study, in which the neurological examination of Parmelee was applied besides a number of other perinatal neurological measures, the Cohen and Parmelee study (1983) could be mentioned. This study aimed to predict intellectual functioning of preterm infants considered at risk. In addition to the neurological exam attentional, manipulative, and exploratory behavior was assessed in the period immediately after birth and four months later. The outcome measures were taken at the age of 5 years. They consisted, among others, of the Bayley Exam and Stanford-Binet test. The results of this study are very modest as far as prediction of cognitive functioning from the earliest measures (neurological, neurophysiological) is concerned. That is, correlations of .30 or less were found. These results, however, are similar to those of other studies (see for a review Drillien, 1972; Saint-Anne Dargassies, 1972, 1977; Sameroff, 1981).

2.2.3. Summary and discussion

Generally neurological investigations result in a high number of false positives, that is, many infants who at an early age present signs of neurological involvement show a favorable outcome. The proportion of false negatives is quite small.

Neurological investigations predict medical outcome measures and may thus also contribute to prediction of cognitive functioning when the investigation is very extensive like the Prechtl exam, or is repeated. An increase of prediction accuracy may also be obtained when neurological assessments are combined with others, for example, psychological tests and neurophysiological measurements, or social class (Dubowitz, Dubowitz, Palmer, Miller, Fawer, & Levene, 1984; Eldredge & Salamy, 1988; Evers-Embden, & Scholte, 1983; Gold, 1979; Rubin & Balow, 1979). Another possible improvement would be the application of more stringent guidelines and development of more definite norms to interpret the scores (Prechtl, 1982). On the one hand neurological abnormality (and associated cognitive development) at later ages has been found to be related to the neurological condition at an early age. On the other hand, however, many neurologically abnormal newborns appear to recover and show no problems at a later age. This underlines a possible weakness in the neurological exams. That is, the neonatal age at which the assessments are performed could be considered a period of instability also for normals. Reliability and validity of the procedures still seem to be in need of extensive investigation, although recently, in a study of Molfese and Thomson (1985), several perinatal scales were compared as to their predictive value and similarities across them. The results of this study show rather low prediction of behavior in later infancy (Bayley Scales obtained at age 6 months). That is, accounted variances of 27% or less were found. Thus the usefulness of the neurological measurements for prediction of later cognitive development for individuals seems, as yet, to be very limited.

3. Development of new and elaboration of related assessment strategies

3.1. General outline

A growing interest in the search for new infant assessment procedures is shown from the late 1960s onward. Considerable research on the application and practical usefulness of these procedures has only recently been taken up, and consequently their use for clinical practice is at present still modest. Moreover, the transfer of findings and techniques from basic and applied research into the hands of the clinician is widely considered a problem and challenge (Fagan, 1988).

Several conditions seemed favorable for the development of new assessment strategies. The first and most relevant was the need for better assessment instruments for young infants, normal and abnormal. Particularly for the latter (i.e., motorically and sensory handicapped, as well as brain-damaged, retarded or premature infants), hardly any assessment instrument or procedure existed. More specifically, most infant tests were not developed for or standardized on abnormal populations (Goldberg & Kearsley, 1983; Rose, 1981; Ross & Leavitt, 1976; Zelazo, 1982). These instruments were also required as a means for indication and guidance of early interventions as well as the evaluation of intervention effects, in specific (e.g., language skills) and broad (e.g., general intelligence) areas of developmental abilities. The interventions could be either medical and/or psycho-educational (cf. Bricker, 1982, and Tjossem, 1976). A second condition was the awareness that young children's abilities were much larger in number and much more advanced than thought before, especially in the area of perception, cognition, and learning (cf. Bower, 1989, and Gottlieb & Krasnegor, 1985). Related to these points was the shift of the investigators' interest from the products of infant abilities to the way these products were achieved, that is, the processes and strategies used (Caron et al., 1983; Field, Huston, Quay, Troll & Finley, 1982; Flavell, 1985; Lipsitt, 1982; Mussen, 1983; Osofsky, 1979; Prechtl, 1982; Rovee-Collier & Lipsitt, 1982). Yet, no clear ways to measure them were readily available. A third condition was the development of new statistical procedures, research design, and technical outfit (Goldberg & Kearsley, 1983; Gottlieb & Krasnegor, 1985; Nesselroade & Baltes, 1979; Porges, 1979).

The assessment procedures presented hereafter require an advanced technical outfit, and are, with some exceptions (e.g., EEG), relatively recent or show a changed view on the infant. Thus, in this section neurophysiological, environmental, conditioned learning, and attention measurements will be discussed.

3.2. Neurophysiological measurements

3.2.1. Outline

Electroencephalography (EEG), sensory evoked potentials (EP, BSER, ERP), ultrasound brain scanning, and computerized tomography scanning (CT) are now established methods for evaluation of brain development, neural intactness and sensory processing of simple stimuli (Bernard, 1985; Ellingson, Dutch, & McIntire, 1974; Gambi, Rossini, Albertini, Sollazzo, Torrioli, & Polidori 1980; Hecox, Cone, & Blaw, 1981; Hrbek, Karlberg, Kjellmer, Olsson, & Rika, 1977; Kotlarek, Zeumer, & Hornchen, 1980; Murakawi, Nakamura, Mizojiri, Aida, & Matsuo, 1981).

Another category of measures is provided by heart rate, crying and sleep-wake patterns, which are thought to reflect CNS functioning, brain maturation or brain damage. The reasons for including these measures in this section are that they are assumed to be related to the processing of external stimuli and to the status or functioning of the neurological system. In addition, they use measurement techniques similar to the neurological procedures.

Generally, all the aforementioned measures are obtained objectively, that is, the responses are recorded automatically. An extensive and advanced technical outfit (including computers) is often needed for the assessments or on-line evaluation of collected data.

3.2.2. Procedures

EEG (not recently developed and related to CT-scanning as well as sensory evoked potentials) and CT-scanning procedures normally are recordings of spontaneous electric brain activity, and brain anatomical status (density of the brain-tissue), respectively. Thus, these methods are supposed to provide information on the status of the CNS, under no specific (experimental) stimulation. These procedures are mainly aimed at guiding medical interventions (Horwitz & Amiel-Tison, 1979). Furthermore, the assumption is, that when applied on very young infants they are useful for diagnosis and short-term prognosis of neurological development, particularly when repeated during the first two years of life. Repetition of measurements is of special relevance when one considers, for example, that early diagnosed damage can be followed by recovery of function of the developing CNS (Karmel et al., 1988).

With regard to EEG, it is generally agreed that the findings are of limited significance as far as prediction is concerned. That is, extremely abnormal EEGs (e.g., an inactive EEG) had a grave prognostic significance, normal EEGs showed good prognosis, while when observing minor abnormalities in the EEG no prognostic significance was found (see review of Dreyfus-Brisac & Ellingson 1977a). For example, Beckwith and Parmelee (1986) studied EEG patterns in newborn preterms and related them with Gesell (at 4, 9, and 24 months) and Stanford-Binet (at 5 and 8 years of age) scores. They found moderate correlations (around .30) of one aspect of the EEG with the aforementioned developmental measures. In addition, they found significant interactions among the EEG measure and measures of the caregiving environment (which had been taken through naturalistic home observations).

In a recent study CT findings were used to predict developmental outcome, measured by the Bayley Scales at 18 months. The subjects of this study were at increased risk for developmental sequelae (asphyxiated infants). The CT scans predicted the outcome for 77% of the total sample correctly, while 10% false positives and 13% false negatives were found (Fitzhardinge, Flodmark, Fitz, & Ashby, 1981). Watanabe and collaborators (Watanabe et al., 1979) combined these procedures (EEG recordings and CT scannings). By combining the procedures increased prognostic value was shown in young subjects with various neurological abnormalities. It was concluded that EEG was a useful adjunct to the CT scan in diagnosis and prognosis. The developmental course of individual subjects or very small groups were described and no group predictions were made.

Ultrasound brain scanning can identify lesions in the brain of very young (preterm) infants. Associations between these lesions and early neurodevelopmental outcome were shown in other studies as reviewed by Costello and collaborators (1988). These authors found highly significant differences in outcome measures at 4 years of age (neurodevelopmental measures and the McCarthy Scale of Children's Abilities) when very preterm infants' scanning results were used for classification.

Event related potentials (or sensory evoked potentials) are transient electrical responses of the CNS to external visual, auditory or somatosensory stimulation (Desmedt, 1977; Dreyfus-Brisac & Ellingson, 1977b; Kurtzberg, Stapells, & Wallace, 1988; Mizrahi & Dorfman, 1980). This category of procedures has proven valuable in assessment of sensory functioning in young infants and difficult-to-test children (Berg & Berg, 1979; Davis, 1985; Despland & Galambos, 1980a, 1980b; Eggermont, 1985; Hecox & Deegan, 1985; Jaffe, 1976; Lester, Karmel, Cantor, & Wheeler, 1983; Salapatek & Nelson, 1985). The procedures also seem to hold promise for assessment of processing capacities. With respect to this latter aspect, the procedures are assumed to provide information on the processing speed and accuracy of sensory stimulation, which may be basic to cognitive development (Lester et al., 1983).

Despite this prospect, some early studies in premature (Dreyfus-Brisac & Ellingson, 1977b) and normal newborns (Henderson & Engel, 1974) did not find significant relations between visual evoked response latencies (VER) and intelligence scores or school achievement in childhood.

Preliminary findings on the prognostic value of auditory brainstem evoked responses (ABRs) for later (at the age of 1 year) cognitive (Griffiths Developmental Scale) and neurobehavioral development point to significant correlations in a sample of high-risk infants. That is, motor delay and neurological abnormalities as well as delayed development in adaptive and reasoning skills were correctly predicted in more than 80% of the infants (Majnemer, Rosenblatt, & Riley, 1988). Murray performed ABRs in high-risk infants and used those measures to predict outcome in Bayley scores, among others. She revealed low predictive power in mental, but higher in motor scores (Murray, 1988a, 1988b). The Majnemer et al. study revealed that the age at which ABRs were taken (not in the first week of life) could be an extremely relevant variable in predictive studies. The latter consideration may have played a role in the modest findings of the Murray studies.

Heart rate patterns (Von Bargen, 1983), crying patterns (Zeskind, 1983), and sleep-wake patterns (Lombroso, 1985; Rose, 1983) are measures of which only preliminary studies have as yet been reported. They are mentioned here, because of their possible future relevance. It has in fact been shown, that individual differences exist in responses to external stimulation and in physiological or biobehavioral responses (heart rate) during different sleep-wake states (Rose, 1983). Sometimes these differences have been related to the infant's maturity level, that is, gestational age (Field, Sostek, Goldberg, & Shuman, 1979) or to specific risk factors, for example, high-risk prematurity and severe brain-damage. In this vein, these measures could be valuable, singly or in conjunction with other measures, in the prediction of development (Bradley-Johnson & Travers, 1979; Karch, Rothe, Jurisch, Heldt-Hildebrand, Luebbesmeier, & Lemburg, 1982; Porges, 1983, 1988).

3.2.3. Summary and discussion

It has been argued that the aforementioned physiological measures are related to observable (future) behavior and thus could have potential to predict later mental abilities. Unfortunately, this assumption has, to date, very limited data support. The rather few studies report similar findings with regard to prediction as achieved with the neurological scales. That is, they show promising prediction of neurodevelopmental short-term outcome, moderate to low predictive power with respect to cognitive development, and a high percentage of false positives. Most studies applied high-risk subjects. The findings of these studies seem applicable in similar populations. For screening purposes, that is, application on a large scale, the procedures do not seem suitable at present. The use of these measurements in combination with behaviorally oriented measurements (such as attention or conditioning) could possibly increase the efficacy of prediction. Moreover, they could provide more information on the conditions and processes playing a role.

3.3. Environmental measurements

3.3.1. Outline

The results of several longitudinal studies have pointed to relations between environmental conditions (e.g., S.E.S., father's or mother's education, adequacy of home environment, ethnicity, family size) and developmental outcome (i.e., IQ, DQ, and school achievement). It has in fact been suggested that these conditions predict outcome better than perinatal measures obtained through medical and neurological assessments, or than behavioral measures obtained through developmental tests. Such suggestions have been made for normal as well as for high-risk children (Beckwith & Cohen, 1984; McCall, 1979; Sameroff, 1981; Sameroff & Chandler, 1975; Smith, Flick, Ferriss, & Sellmann, 1972; Wallace, 1988). However, the predictive value of these environmental conditions has not been evidenced before the age of 1 to 1 1/2 year, presumably because these conditions could not have exerted significant influences in the first few months of life, and, in addition, those influences need time for their manifestation. Most correlations found were below .30. These correlations, however, showed to be stronger when the ages at which outcome measures were taken increased. These findings indicate that the aforementioned conditions contribute to only a small portion of the variance in cognitive outcome (Golden & Birns, 1976; Jackson, 1982; Siegel, 1984).

Although a stream of studies (see for extensive reviews Hunt, 1979a, 1979b; Wachs & Gruen, 1982) have suggested influence of environmental variables on outcome, others (i.e., those on early environmental enrichment) have found no (strong) evidence supporting such claims. In spite of the assumption that environmental conditions play a significant role in developmental outcome, and that optimal conditions improve this outcome, the latter studies proved rather disappointing with regard to long-term as well as, occasionally, short-term effects (Golden & Birns, 1976). The data of these studies were taken by some authors (geneticists, i.e. Jensen, 1969) as additional evidence that intelligence is to a great extent genetically determined and therefore cannot easily be influenced by environmental variables. By contrast, other authors (nurture oriented) interpreted the results of the aforementioned studies in quite a different vein. That is, they argued that the environmental variables included in such intervention programs were too broadly defined with regard to both their characteristics and the extent of their manipulations. Moreover, they maintained that the instruments adopted to measure the effects of the programs were not suitable to discriminate subtle though significant gains in development and behavior. In addition, the effects of the programs might for part of the subjects have been beneficial, but for others negligible or negative, thus showing the overall effects being small or not present (Bradley & Caldwell, 1978; Horowitz & Paden, 1975; Wachs, 1984; Walberg & Marjoribanks, 1976).

On the basis of the second hypothesis new research interests have developed concerning the definition of environmental variables which could more likely be rated as "real precursors" of later development (i.e., aspects of the physical and social environment). At the same time efforts have been made to quantify those variables, as to their frequency of occurrence, duration and intensity, and contingency value, among others (Wachs & Gruen, 1982). In addition, initiatives have been taken with regard to assessment procedures which could prove more suitable and sensitive in measuring any change in developmental and behavioral patterns (Ross & Leavitt, 1976; Wachs & Gruen, 1982). Those research efforts, however, did not discount the importance and possible influence of biological and medical variables or conditions (see for example, Gewirtz and Petrovich, 1982; Greenberg and Crnic, 1988; Ramey and Baker-Ward, 1982; Stratton, 1982a). In fact, several studies evaluated their differential impact on and interaction with environmental variables (Bakeman & Brown, 1980; Bradley & Caldwell, 1980; Cohen & Parmelee, 1983). In the following section some procedures and studies aimed at investigation of environmental variables that could have relevant effects on developmental outcome are discussed.

3.3.2. Procedures

Barnard and collaborators (Barnard, Bee, & Hammond, 1984), Beckwith and Cohen (1984), and Bradley and Caldwell (1982, 1984a) have investigated the quality of the caretaking environment. They have analyzed it in terms of (a) emotional and verbal responsivity of the mother, (b) avoidance of restriction and punishment, (c) organization of physical and temporal environment, (d) provision of appropriate play materials, (e) maternal involvement with children, and (f) opportunities for variety in daily stimulation. In order to quantify these various aspects the Caldwell Home Inventory (Bradley & Caldwell, 1978) has been developed. HOME (Home Observation for Measurement of the Environment) scores obtained on normal infants in the first year of life showed significant relationships with later scores on developmental and intelligence tests, that is, the Bayley and the Stanford-Binet. Correlations of around .50 of the total HOME score with later IQ were found. In addition, the subscale Play Materials played a significant role in multiple relationships (Bradley & Caldwell, 1976, 1980, 1984a, 1984b; Ramey, Farran, & Campbell, 1979). A follow-up by Bradley, Caldwell, and Rock (1988) at the age of 10-11 years and applying the Science Research Associates (SRA) achievement test battery, showed only minor relationships with the home environment at 6 months of age (correlations smaller than .20). In a collaborative study, applying the HOME measure at 12 months, the Bayley Mental Development Index at 24 months and the Stanford Binet IQ at 36 months of age, Bradley et al. (1989) found correlations of about .50 among the aforementioned measures. It could be concluded from these studies that higher correlations are found when HOME measures are taken at later age. Particularly from the age of 1 year on, the correlations with outcome measures seem substantial. Furthermore, HOME gives clues as to which aspects of the environment play a significant role in development. How these aspects affect cognition has not been investigated.

A number of authors (Carlson & Bricker, 1982; Pianta, Sroufe, & Egeland, 1989; Vietze, Abernathy, Ashe and Faulstich, 1978; Watson, 1976; Watson & Ramey, 1972) have investigated the contingency value of environmental stimulation and caretaker competence, which they considered of fundamental importance for infant development.

The caretaking environment was analyzed in terms of patterns of infant and caretaker's behavior (Coates & Lewis, 1984; Greenberg & Crnic, 1988; Pianta et al., 1989; Vietze et al, 1978), infant's contingency perception (Watson, 1979), social and physical environmental contingencies (Carlson & Bricker, 1982), and infant's capability to discover contingent relationships between its own acts and their effects on the environment (Riksen-Walraven, 1978). Generally, observations of caregiver-infant interactions (i.e., occurrence of temporal distribution of stimuli and responses) provide for quantification of the aforementioned variables (see for a discussion on methodology also Yarrow & Anderson, 1979). Watson (1979) also suggests assessment of learning rate through classical or operant conditioning as a measure of perception of contingency in the environment.

That the contingency value of the environment could be an important aspect in development has been suggested in a study of Coates and Lewis (1984). They found significant single (most correlations around .30) and multiple (multiple Rs around .60) relationships between measures of maternal interaction (at 3 months of age) and later developmental status. The latter status was measured by several cognitive developmental tests, at the ages of 12 and 24 months, and 6 years.

More recently another aspect has been stressed, that is, the role the child or the child's characteristics play within the developing caretaking environment. Ultimately this could prove to play an essential role in long-term cognitive development. On this latter aspect no clear empirical evidence is available. However, at least for atypical infants (prematures) there is some support with regard to the influence of differential interactions (as they seem to exist) between infant and primary caregiver on cognitive outcome (Cohen & Beckwith, 1979; Cohen & Parmelee, 1983; Klein, 1984; Ramey et al, 1979). For example, in the Cohen and Parmelee (1983) study caregiver-infant interaction measures were taken with eight-months-old preterm infants. The relation between the latter measures and nine- and 24-months Gesell and five-year Stanford-Binet scores showed to be significant but modest (correlations of .45, .35, and .28, respectively). In this study the interaction variable together with years of maternal education represented the best predictors. Moreover, they found that measures of caregiver-infant interaction distinguished between infants whose cognitive performance improved and those who continued to show delays.

3.3.3. Summary and discussion

In conclusion, the idea that early experience (particularly in the preschool years) is of special importance for psychological development is now not in dispute. Consequently, a relationship between early environmental conditions and later development should exist. This relationship could be age-dependent. Which parameters of the environment will promote an effect on later development and which variables would be most sensitive to reveal this relationship at an early age, is, to a certain extent, still an open question. Moreover, the reliable measurement of these variables seems not yet available, perhaps with the exception of the HOME measurement which has been and is being studied extensively at the moment.

3.4. Attention measurements

3.4.1. Outline

Attention is one of the most widely studied measures in relation with cognitive development in young children. It has been proposed as very relevant for assessment of cognitive and sensory abilities and for prediction of later cognitive performance (Cohen, 1981; Gottlieb & Krasnegor, 1985; Miller, Ryan, Aberger, McGuire, Short, & Kenny, 1979; O'Connor, 1980; O'Connor, Cohen, & Parmelee, 1984; Rose, 1983; Teller, 1979).

Measures of attention are considered to reflect the child's proficiency to process information (Friedman, Jacobs, & Werthmann, 1981; Gottlieb & Krasnegor, 1985). The ability to control attention is assumed to be a prerequisite for memory and learning (Bower & Hilgard, 1981; Finkelstein, Gallagher, & Farran, 1980; Flavell, 1985; Kail, 1984; Moscovitch, 1984; Olson & Strauss, 1984). Moreover, some authors maintain that strategies can be developed for assessing these areas, particularly for special populations, such as high-risk, motorically or mentally handicapped infants (Berg & Berg, 1979; Cohen, 1981; Finkelstein et al, 1980; Friedman, S.L. et al, 1981; Gardner & Karmel, 1983; Hayes, Ewy, & Watson, 1982; Lewis & Brooks-Gunn, 1981; Miller et al, 1979; Sigman, Cohen, & Forsythe, 1981; Stratton, 1982a). Different procedures are used to investigate attentive processes. They are discussed in the following section.

3.4.2. Procedures

Several approaches have been used to study attention in young children. These are a) the habituation-dishabituation technique, b) the paired comparison paradigm, c) the visual preference technique, and d) sustained attention.

The first two procedures are both called novelty preference techniques (Sophian, 1980). In these procedures an important phenomenon is habituation. Habituation has been defined as a decrement in responding with repeated stimulus presentation or decrease in attention due to prolonged stimulus presentation (Stephenson & Siddle, 1983; Tighe & Leaton, 1976). Habituation is considered an elementary form of learning (Bower & Hilgard, 1981; Siddle, 1983). For habituation to occur, a subject has to attend to the stimulus, to process and to store it (and to discriminate it from another discrepant stimulus: dishabituation). The habituation-dishabituation technique consists of repeated presentations of the same stimulus followed by a new, discrepant stimulus. The paired-comparison paradigm is a variation of the previous technique. It also includes two steps, that is, a prolonged presentation of one stimulus (i.e., 1-2 min) followed, after a delay, by a simultaneous presentation of the recently exposed and a novel stimulus. In the visual preference technique the subject is expected to show a preference for one over the other simultaneously presented stimulus. By sustained attention, the duration of the subject's attention to a novel stimulus is observed.

Quantification of the aforementioned variables is obtained by observing changes in looking responses, sucking rate, heart rate, or the total amount of time a child looks at a stimulus.

Among the studies applying the novelty preference techniques the following may be cited: Bornstein and Benasich, 1986; Fagan and McGrath, 1981; Fagan and Singer, 1983; Lewis and Brooks-Gunn, 1981; Lewis and Taft, 1982; Miller, Spiridigliozzi, Ryan, Callan, and McLaughlin, 1980; Moss, Colombo, Mitchell, and Horowitz, 1988; O'Connor et al, 1984; Ruddy and Bornstein, 1982; Werner and Siqueland, 1978.

For example, Fagan and collaborators (Fagan & McGrath, 1981; Fagan & Singer, 1983) report on a longitudinal study with normal children. They applied visual recognition tasks at the age of 4 months and vocabulary tests (PPVT) at the ages of 4 and 7 years. The authors found significant relations between infant recognition memory, based on novelty preferences, and later verbal tests (correlations of .37 and .57, respectively), but not between habituation rate and verbal IQ.

In an interesting investigation of Lewis and Brooks-Gunn (1981), some of the early infant measures were studied as to their predictive power. The authors found, in a group of normal infants, that a) conventional tests of sensorimotor development, administered at 3 months, failed to predict intellectual outcome at 2 years, b) rate of habituation at 3 months had no consistent relation to later intelligence, while c) preference for visual novelty was significantly associated with two-year functioning ($r = .52$ to $.40$), measured with Bayley MDI and Escalona-Corman.

Lewis and Taft (1982) also reported data on the development of visual attention in infants, high-risks and normals, and some anecdotal evidence for the possible significance of this measure for detecting dysfunctions in visual information processing. That is, the high-risk subjects showed no response decrement over the repeated presentations of a stimulus until the age of 1 1/2 to 2 years, while the normals did, in increasing fashion with increasing age. In addition, they showed differences of response decrement existing between samples of normal and high-risk infants in attention to auditory stimuli. This was measured by changes in the sucking rate during a habituation task.

Ruddy and Bornstein (1982) investigated the predictability of cognitive differences at 12 months from infant and maternal behaviors at 4 months (child's age). They found that the (normal) infants who showed more and faster habituation to visual stimuli at 4 months, had higher scores on the Bayley Scales (r = .47) at 12 months.

Comparable results have been reported by Miller et al. (1980). They applied similar tasks and cognitive measures with normal children at the ages of 2-3, 15, 27, 39 and 51 months, respectively. The obtained correlations of the habituation measures with the cognitive measures were quite modest (around .30). Faster habituators at the age of 51 months tended to be somewhat advanced cognitively, compared to slower habituators.

O'Connor et al. (1984) reported on a longitudinal study with full-term and preterm infants. They found a rather high and statistically significant (r = .60) correlation between infant novelty scores (at 4 months), measured by cardiac responses to auditory stimuli, and five-year intelligence scores (Stanford-Binet). In this study mother's education was also a significant predictor of infant outcome. In a previous study (O'Connor, 1980) a significant relationship was found between response to novelty at 4 months and 18-month Bayley Mental scores.

Sucking rate as an operationalization of attention was used by Werner and Siqueland (1978). The results of their investigations on 16 premature infants revealed that a) the experimental subjects responded differentially to novel versus familiar stimuli, and b) in a subgroup of 7 infants individual differences in visual exploration and responsiveness to novelty were negatively correlated with concurrently taken variables, that is, perinatal complications, and positively correlated with maturation level. These correlational (positive and negative) levels were always above .50. From the 36 subjects originally available 20 were rejected because of failure to complete the testing sessions.

With regard to studies involving the visual preference technique the following may be cited: Miranda, 1976; Miranda and Hack, 1979; Sigman and Beckwith, 1980; Sigman et al., 1981; Sigman, 1983. For example, Miranda and Hack showed that differences existed in the developmental course for Down's syndrome and normal infants. In addition, they considered the procedure to be promising for studying the relationships between early visual perception and later intellectual performance. In fact, they provided evidence that individual classifications, based on the visual preference technique applied in high-risks, could discriminate quite successfully normal from abnormal developing infants (Miranda & Hack, 1979). However, the procedure is not yet standardized and behavioral state of infants seems to be a major problem (see also Gardner & Karmel, 1983).

The application of selective attention measures for long-term prediction has also been pursued by Sigman and collaborators. In a group of preterm infants those authors found significant (but low) correlations between a visual attention risk score, attention duration (e.g., visual fixation of a single stimulus), both at term age, and intelligence (Stanford-Binet) at 5 years (r = -.29 and -.25, respectively). However, a visual preference risk score at 4 months (based on preferences for specific stimuli, e.g., preference for photographed faces rather than geometric forms) did not correlate with IQ at 5 years. Similar relationships were found, applying the same attention measures and Bayley and Gesell scores at 18 and 24 months, respectively.

The sustained attention measure has been applied by Kopp and collaborators (Kopp & McCall, 1982; Kopp & Vaughn, 1982). In a task, demanding exploratory manipulation of colored objects, sustained attention was used with eight-months-old babies who were considered at risk because of their premature birth. This measure, applied together with SES, gestational age and others, contributed significantly to the prediction of status on the Bayley Mental Scales, applied at 2 years of age. Simple correlations among the aforementioned attention and outcome measure, however, revealed only modest relationships ($r \leq .21$).

3.4.3. Summary and discussion

Regarding the attention measurements, it could be concluded that the different authors apply different tasks with regard to stimuli and difficulty. Promising results have been reported, both regarding predictive as well as concurrent validity of attention measures, thus making these measures of potential usefulness as a cognitive index (Bornstein & Benasich, 1986). The reported correlations among early attention and later outcome measures varied between .25 and .60. In addition, in some studies the correlations among attention measures (applied at the first six months of life) and IQ showed at older ages to be higher than at younger ages. It is not clear whether the results of the studies depended on the age at which the measures (predictor as well as outcome) were taken, the quantification of the attention measures (e.g., fixation frequency or duration, in total or only of the first fixation occasion, and nature of measurement of the responses), or the characteristics of the populations studied, among others. In this respect, it would be interesting to find out, what the intercorrelations between different measures of habituation and attention are (McCall, 1981). One of the problems noticed with the attention measurements is that the state of the infants, which has consequences for the reliability of the measurements and possible subject loss (Gardner & Karmel, 1983). In addition, the question has been raised whether an infant who examines stimuli for long periods of time is analyzing the world carefully or is slower to take in visual input or regulate his or her own attention processes (Sigman, 1983). Another question concerns the relation between individual differences in infant attention patterns and environmental elicitors and responses or experience with environmental contingencies. Moreover, it is not yet clear which stimulus types, response measures, and study-design would be most suited with regard to the research questions. Related to the latter issue is the notion that the attentional behaviors involved in the acquisition of information may be different from those involved in subsequent recognition memory and may therefore tap different aspects of information processing (Rose, Feldman, McCarton, & Wolfson, 1988; Rose, Feldman, Wallace, & McCarton, 1989).

3.5. Conditioned learning assessments

3.5.1. Outline

Investigation of learning and memory in animals and human beings has a long tradition in psychology. Generally, this line of research has concentrated on investigation of learning capacities and processes, mostly in organisms who already possessed relatively mature perception and response systems (Friedman, Das, & O'Connor, 1981; Hall, 1982; Spear & Campbell, 1979). While the behavioral repertoire of the human adult is enormously sophisticated, as has been illustrated in numerous studies, but also in everyday life, that of the neonate is rather restricted. During development the behavioral repertoire extends in quality and quantity through processes of maturation and learning. However, the study of such learning per se, as well as learning with regard to assessment of developmental potentials (as distinct from already acquired abilities or skills), has only recently been stressed. For example, Rose (1981) and Rovee-Collier and Lipsitt (1982) underscored the need for measures assessing infants' learning or memory. More in particular, they maintained that those measures should not be the kind of global indexes which developmental scales were assumed to give. Moreover, Watson (1976) argued that infant learning might provide useful predictive relationship for future individual differences in intelligence. However, his main conclusion was that more empiric evidence was needed on the relation of learning and intelligence. Knowledge about this relation in infancy (but also in child- and adulthood) was nearly absent. In addition, Lipsitt (1979) contended that assessment of the learning ability of infants could have the potential to provide clues as to which infants possess the requisites for successfully adaptation to their environment. In particular, this latter issue seems of utmost importance for infants who are born with birth defects or undergo early risks such as prematurity, anoxia, obstetrical medication, and other neonatal hazards.

For human infants much research on learning, besides the already discussed habituation procedures, has been carried out through classical and operant conditioning procedures. These conditioning paradigms were thought to be remarkably appropriate to infant subjects (Olson & Sherman, 1983; Ross, 1966; Ross, Headrick, & MacKay, 1967; Ross & Ross, 1973; Ross & Leavitt, 1976). Of the two conditioning methods, the operant procedures were most widely applied (Fitzgerald & Brackbill, 1976; Lancioni, 1978).

Although at the beginning of the century a few studies had appeared (Marquis, 1931; Myers, 1908; Valentine, 1914; Watson & Raynor, 1920), research on infant conditioning has built consistency and variety during the 1960s and the 1970s. The presentation of numerous reports in this period, and thereafter, indicates a heightened interest in the field of infant behavior and learning. The change has been ascribed to a revival of the interest in and a reevaluation of the nature-nurture issue, innovations in instrumentation and research-methodology, and the popularity of behaviorism. Human conditioning is now concerned with processes involved in learning and how this learning is translated into the behavior observed. In this way conditioning can be viewed as a means to change behavior, as an adaptive learning process, and, very importantly, can be used as a tool for studying all kinds of psychological processes (cf. Davey, 1981, 1987; Fitzgerald & Brackbill, 1976; Fitzgerald & Porges, 1971; Lancioni, 1978; Lancioni, 1980; Lipsitt, 1969; Rovee-Collier & Lipsitt, 1982).

Different kinds of classical and operant conditioning procedures, as employed in child studies, will be outlined. Several studies will be mentioned, applying stimuli, responses or paradigms relevant for the study of infant abilities.

3.5.2. Procedures

3.5.2.1. CLASSICAL CONDITIONING

Classical conditioning procedures may vary in several respects, but in essence they consist of the transfer of stimulus control from a stimulus that reliably elicits a response (unconditioned stimulus or UCS) to one that was previously neutral, the conditioned stimulus or CS. The establishment of a CS-UCS contingency results in a so-called conditioned response or CR (Dawson & Schell, 1987; Lancioni, 1978; Olson & Sherman, 1983).

Two kinds of procedures may be applied, that is, simple and complex procedures. In simple procedures, applying only one conditional stimulus (CS), four paradigms can be distinguished, that is, simultaneous, delayed, trace and temporal. This distinction is based on the temporal relationship between the CS and UCS. In complex conditioning procedures two paradigms can be implemented, that is, differential conditioning applying two conditional stimuli, and stereotype conditioning. In the latter strategy a sequence of behavior is established through the presentation of a constantly fixed series of conditional stimuli. Differential conditioning involves two CSs, only one of which is associated to the UCS.

Classical conditioning procedures were successfully applied with newborns (Blass, Ganchrow, & Steiner, 1984), young infants (Ingram & Fitzgerald, 1974; Lancioni & Hoogland, 1980; Lancioni et al., 1985; Lintz, Fitzgerald, & Brackbill, 1967; Little, Lipsitt, & Rovee-Collier, 1984), young normal children (Irzhanskaia & Felberbaum, 1967), and mentally subnormal children (Franks & Franks, 1962; Grings, Lockhart, & Dameron, 1962; Lancioni, Coninx, & Smeets, 1989; Ross et al, 1967). Successful conditioning implies the reliable establishment of conditioned responses (CRs) upon the presentation of conditional stimuli.

Conditional stimuli (CSs) used with the procedures were *temporal*, that is, intertrial intervals of fixed length (Brackbill & Fitzgerald, 1972; Brackbill, Fitzgerald, & Lintz, 1967; Fitzgerald, Lintz, Brackbill, & Adams, 1967), *tactile*, for example, palm press and forehead stroking (Blass et al., 1984; Connolly & Stratton, 1969), *auditory*, for example, pure tones (Franks & Franks, 1962; Lancioni & Hoogland, 1980; Lancioni et al., 1985, 1989; Little et al., 1984; Naito & Lipsitt, 1969), *visual*, such as light flashes (Crowell, Blurton, Kobayashi, McFarland, & Yang, 1976), and *olfactory*, for instance, odor of mint (Irzhanskaia & Felberbaum, 1967).

The unconditioned stimuli were a) desirable or appetitive (i.e., sucrose solution), which would elicit approach reactions (Blass et al., 1984), and b), undesirable or aversive (i.e., a stream of air on the eye), which would presumably elicit avoidance reactions (Little et al., 1984).

The nature and intensity of the conditioned response (CR) is directly determined by the choice of the unconditioned stimulus which should elicit it. Conditioned responses can be divided into *autonomic* (i.e., heart rate, Clifton, 1974), and *somatic*, (i.e. eye blinking and sucking, Blass et al., 1984; Lintz et al., 1967).

It has long been a question whether in newborns conditioning could take place. In fact, sufficient evidence now suggests a positive answer (Blass et al., 1984; Fitzgerald & Brackbill, 1976; Lintz et al., 1967; Little et al., 1984). Another and related question concerns the neurological structures necessary or sufficient for conditioning to take place and, in addition, the nature of the conditioning process when different neurological structures are involved. This question is still open to investigation. Individual differences in conditionability seem to exist (Maltzmann & Mandell, 1968; Sokolov, 1963), but if these differences remain stable over time and are indicative of different behavior and development (and thus can be used to predict later learning performance), is unclear since too few studies have addressed such issues.

3.5.2.2. OPERANT CONDITIONING

Operant conditioning consists of the occurrence of reinforcement (stimulus), dependent on the presentation by the subject of the response to be learned. Thus, the probability of occurrence of this response is affected, that is, increased or decreased. Two different procedures may be distinguished, that is, free-operant and discrimination learning (Lancioni, 1980; Reese & Lipsitt, 1970; Sameroff & Cavanaugh, 1979).

In free-operant conditioning the subject is free to perform the target response as frequently as possible within a predetermined time interval, usually called the conditioning session. During this time the target response is reinforced contingently, even though not necessarily on a continuous schedule. This method does not provide the organism with external cues to signal the availability of the reinforcement upon the emission of the response, except for the occasional presence of the experimenter, equipment or response manipulandum. In discrimination learning, the subject is generally reinforced for performing the target response to one stimulus and non-reinforced or punished for responding to other stimuli. Several different procedures are used. All procedures apply one or more specific signals (cues or discriminative stimuli: S^Ds) to show the subject the availability of a reinforcement occasion or a fixed time-period in which reinforcement is provided.

Operant conditioning procedures have been successfully applied to infants and children, both normal and abnormal (Fitzgerald & Porges, 1971; Lancioni, 1978, 1980; Lipsitt, 1969; Olson & Sherman, 1983).

Among the studies that have used operant conditioning for assessment of individual or group differences in normal or handicapped populations, the following would seem more relevant : Gekoski, Fagen and Pearlman (1984) and Millar (1976).

Gekoski and collaborators (1984) applied a mobile conjugate reinforcement paradigm in young preterm and fullterm infants (slightly older than 50 weeks conceptional age). They reported differences in acquisition rate and retention (one week) of an operant response (footkicking) between groups, with the preterm subjects being disadvantaged. That is, the full-terms acquired the task in a single session, while the preterms needed two sessions. In addition, the retention session showed higher relative response rates for the full-terms than the preterms. Whether the preterm's retardation would persist and/or generalize to other types of learning situations, remained unanswered.

Millar (1976) reported differences in response acquisition between six- and nine-months-old normal infants. He applied contingent social reinforcement of a manipulative response (a non-social response). The frequency of response was entered in analyses of variance between as well as within the two groups with age, contingency and experimental phase (baseline, contingency, and extinction) as factors. The findings of the study, that is, reliable acquisition and extinction effects in the older but not in the younger group, point to two important issues. First, there seem to be age differences in the capability of reaching reliable response acquisition. Secondly, this phenomenon might be related to stimulus or response characteristics or to their relationships, since successful operant conditioning was previously demonstrated in younger subjects.

3.5.3. Summary and discussion

In summary, many authors maintain that learning processes play an important role in cognitive development (Lipsitt, 1982; Reese & Lipsitt, 1970; Sameroff & Cavanaugh, 1979). At the same time they suggest that a) assessment of learning ability at a very young age, but also at later ages, could provide very useful information as to assessment of developmental or learning potential (Gekoski & Fagen, 1984; Hamers & Ruijssenaars, 1984; Lipsitt, 1979; Millar, 1976; Olson & Sherman, 1983; Watson, 1976), b) learning processes in infants and young children can be studied through classical and operant learning procedures, and c) assessment through conditioning procedures would in all probability provide useful information as to early intervention procedures (Lancioni, 1980). However, only a very limited number of studies have applied conditioning procedures for the assessment of learning rate and have related the acquired learning rates to subsequent psychological development. Consequently, concrete information with regard to optimally suited procedures, stimuli and responses as well as ages for such assessments is lacking.

In view of the aforementioned considerations studies are needed to investigate the possibility of application or, in case of necessity, adaptation of procedures directed at the aforementioned goals. In fact, some evidence points towards the practical feasibility of such adaptations (Lancioni & Hoogland, 1980; Lancioni et al., 1985). Lancioni and cooperators succesfully applied classical conditioning procedures for the assessment of hearing abilities in young normal and abnormal infants. The applied procedures required a rather small number of training trials. Moreover, operant procedures have also proven to be feasible at rather young ages (Gekoski et al., 1984). A further requirement of learning assessments for prediction of subsequent performance could be the following. The assessments should preferably consist of more than one and different items or learning tasks, both in order to prevent extraordinary scores due to incidental state problems (affecting performance on a task) and to encompass a wide range of stimuli and responses.

4. Assessment of learning rate: An experiment

4.1. Outline

In the previous chapters, procedures considered relevant for assessment of individual differences and their development were reviewed. It was outlined that refinement and extension of traditional assessment procedures had not resulted in a substantial increase in prediction accuracy with respect to subsequent development. That is, significant correlations were found between early and later development, but these were generally too low to be used for diagnostic purposes. New procedures and measures (others than the aforementioned) have been considered to have potential as assessment means, although sufficient data are not available yet.

One of these measures, recently emphasized, is the learning ability, and more specifically conditioned learning (Lipsitt, 1979; Rose, 1981; Rovee-Collier & Lipsitt, 1982; Watson, 1976). Many studies have shown that conditioning is possible with infants of different ages (for reviews see Lancioni, 1978, 1980). However, these studies were generally not aimed at studying individual learning rates. Moreover, the relationships between learning rate and developmental outcome have not been investigated (Lancioni, 1978, 1980; Lancioni et al., 1985).

Alternative measures, also assessing cognitive processes (e.g., through habituation techniques), have been found useful in assessment of high-risk and handicapped children. Assessment of learning through conditioning procedures seems warranted.

Therefore, the present study was aimed at assessment of individual differences in learning ability as measured by classical and operant conditioning procedures. In addition, the stability of these differences was investigated by the application of a longitudinal design (Hunt, 1979a; Porges, 1979). Furthermore, the efficacy of some predictive models based on variables from the aforementioned procedures and outcome assessments was explored.

Two groups of subjects participated. One group was made up of normal infants, the other group of high-risk and handicapped infants.

The first assessment was carried out when the subjects were 4 months of age. This age was corrected for prematurity (Barrera, Rosenbaum, & Cunningham, 1987; Siegel, 1983; Wilson, 1987). This age was chosen on several grounds: a) it avoids the instability of the postnatal adjustment (Brazelton, 1973; Knobloch et al., 1980; Prechtl & O'Brien, 1982), b) the relationship between infant and environment is much more structured at 4 months than during the neonatal period, c) sensory, neuro-motor, and learning rate assessments are easier to perform and more reliable, d) the investigation of the learning rate through conditioning paradigms requires a relatively short time (Lancioni & Hoogland, 1980; Lancioni et al., 1985; Papousek & Papousek, 1982), and e) it seems still early enough to start intervention if this were necessary. The first assessment consisted of classical conditioning procedures.

The second assessment took place when the infants were 6 months of age. This was thought to be an appropriate age for implementation of the operant conditioning procedures, because the infants are old enough to be able to learn a contingency and to perform the required response, that is, head-turning (Finkelstein & Ramey, 1977; Lancioni, 1978; Reese & Lipsitt, 1970).

The last assessment (including an operant conditioning procedure and a Bayley Exam) was scheduled at the age of 18 months because that age is considered a milestone (Illingworth, 1975, 1983). That is, developmental measures obtained at that age are rated as representative of the condition of the subject and subsequent performance (McCall, 1979). Moreover, the Bayley Scales are generally considered a useful assessment tool for measurement of mental abilities, also in deviant populations, particularly at this age (Ross, 1985).

In addition to the aforementioned assessments at the age of 6 months home-observations were planned in order to provide some additional information on the way infants engage with the surrounding world and to explore possible relationships with the other measures applied (Stratton, 1982b).

4.2. Method

4.2.1. Subjects and setting

Two groups of subjects participated in the study, that is, a group of high-risk and handicapped infants and a group of normal infants. General characteristics of both groups are described in Table 1. For neither group of infants social class was a selection criterion and thus the subjects were from various social backgrounds. The subjects were recruited in the period 1982-1984. All the parents agreed to cooperate and no payment was provided.

4.2.1.1. HIGH RISK AND HANDICAPPED INFANTS

Thirty-nine high-risk and handicapped infants participated in the study. A high-risk infant could be defined as one who is at greater than average risk for later deviances in behavior because of membership in some identifiable population (Sameroff & Seifer, 1983). Of this group one subject did not complete the study because of death. Two subjects were not able to attend to the complete testing-procedures because of sensory impairment (one subject had hearing deficits and the other had visual problems). The aforementioned three subjects were eliminated from the study sample. On six subjects incomplete data were obtained. However, these latter subjects were maintained in the study sample. Of these six subjects, a) three subjects did not attend to the six months operant conditioning assessments because of equipment problems, b) another subject completed the study except the assessment at 18 months because of motoric disability (e.g., hypotonicity), and finally, d) two subjects with leg-paralyses were not able to complete the Bayley Motor Scale. In conclusion, data on 36 subjects could be used. Complete data were available on 29 of these 36 subjects.

These subjects had been recruited from patient-records from the Sint Radboud University Hospital and the Canisius-Wilhelmina Hospital, Nijmegen, Holland. Selection of the subjects had been carried out by pediatricians of the same hospitals from the population of patients as occurred in the study-period. The pediatricians were normally in charge of the medical treatment and follow-up and were informed about the general experimental procedures. Selection of the subjects took place according to several criteria. The first was the developmental vulnerability, as indicated by a variety of medical risk factors (e.g., prematurity, metabolic and infectious diseases and other disorders with possible CNS involvement). The second criterion was the medical status of the subjects. That is, they had to be in a stable condition. Thus, drastic changes in their general behavior were not to be expected. A final and obvious criterion was the behavioral status of the children. That is, they had to be responsive to external stimulation. This criterion would seclude children with severe sensory impairment and extreme neuro-behavioral deficits.

For infants meeting the aforementioned criteria a recruitment procedure was initiated. The pertaining pediatrician, who had regular contacts with the parents, explained about the nature of the experiment and asked if they were willing to let their child(ren) participate. In case of a provisionally positive answer, the experimenter contacted the parents and provided detailed information. Hereafter the parents made the definite decision with regard to consent. General characteristics of the study sample are described in Table 1. Additional characteristics are as follows.

Table 1
General Characteristics of the High-risk and Handicapped Group (High-risks) and the Normal Group (Normals)

	High-risks(n=36)	Normals(n=12)
Gestational age (range in weeks)	26 - 41	39 - 41
Birthweight (range in grams)	625 - 3200	2800 - 3700
Males	18	6
Females	18	6

As shown in Table 1 gestational age as well as birthweight of the children varied widely. This wide range was due to the fact that preterm as well as full-term babies participated. Of all these children 22 were born with a gestational age of 37 weeks or less (Sweet, 1979). The latter group of children included appropriate- as well as small-for-gestational-age infants and/or infants with associated problems. In addition, one child, which was born preterm, was affected with Down's syndrome. Two children had spina bifida in association with other neural tube defects. Another three subjects had hydrocephalus. Neonatal CNS infections were the reasons for inclusion in the study for two infants.

Another two children were included because they had shown a history of apneic episodes or other respiratory problems (possible near-miss sudden death). Of the remaining four children one had a metabolic disease, one feeding problems with associated disturbance of growth, another infant showed seizures, while the last subject had disturbances in the functioning of the kidney(s). All subjects were hospitalized for prolonged periods after birth, but not during the study period.

In the tables as well as in the text this group of subjects will be referred to as the high-risk (and handicapped) group.

4.2.1.2. NORMAL INFANTS

The initial sample consisted of 12 subjects, 6 females and 6 males. Three subjects underwent only part of the investigations because of change of residence or other family circumstances. They were born with normal gestational age and birthweight and were clinically normal, that is, they showed no signs of disorder or major disease. These subjects had been selected either randomly from birth-records from the Department of Obstetrics of the Sint Radboud Hospital or by directly contacting the parents. In the first case the recruitment procedure was similar to the one used for the high-risk and handicapped group. In the second case the parents were acquaintances of the experimenter (4 infants).

4.2.1.3. SETTINGS

All the testing took place in a quiet room at home. The observations were done in the room where the infant normally was during awake periods (generally the living-room). The reason why the experiment was carried out at home was that parents were more willing to cooperate. Moreover, subjects as well as the parent(s) were more comfortable in their own natural environment. This would eventually prevent unnecessary restlessness in the infant, and would hopefully assure normal behavior in subjects and caretakers.

4.2.2. General procedural conditions and design

Throughout the study several assessments were scheduled at fixed ages and time intervals. All the subjects underwent the same investigations at the same ages. An outline of the learning rate assessments is presented in Table 2.

The experimental investigations took place when the subjects were 4, 6 and 18 months of age. Moreover, at the age of 6 months home-observations were carried out. The first assessment at the age of 4 months consisted of three classical conditioning procedures. The second assessment, at the age of 6 months, included two operant conditioning procedures. The third and last assessment took place when the subjects were 18 months old. It comprised an operant conditioning procedure and the administration of the Bayley Mental and Motor Scales.

Table 2
Learning Rate Assessments

4 Months.	Three classical conditioning procedures:	
	Conditioned Stimulus:	a) Auditory b) Visual c) Tactile
	Unconditioned Stimulus:	Air-puff
	Response:	Defensive reaction
6 Months.	Two Operant Conditioning Procedures:	
	Discriminative Stimulus:	a) Visual b) Auditory
	Reinforcers:	Animated Toys
	Response:	Head Orientation
18 Months.	One Operant Conditioning Procedure:	
	Discriminative Stimulus:	Visual
	Reinforcers:	Animated toys
	Response:	Leverpulling

4.2.3. Measures and materials

4.2.3.1. ASSESSMENTS AT 4 MONTHS

When the children were 4 months old the first learning rate measures were taken. These measures were obtained by classical conditioning procedures (relying on auditory, visual and tactile CSs). The learning rate was defined as the rate at which the subject acquired a conditioned response, that is, the number of training trials a child needed to reach a learning criterion.

The apparatus consisted of an electronic control device connected to an audiometer and to a cylinder filled with compressed air. The electronic device was a general-purpose system which could control a) the onset and duration of a tone (CS) derived from the audiometer (or cue on the onset and duration of visual and tactile CSs), b) the onset, duration, and intensity of the airpuff (UCS) obtained from the air-cylinder and emitted through a plastic tube, and c) the interstimulus and intertrial intervals. In addition it provided a continuous visual display of the number of trials passed. The audiometer was a Grason Stadler Manual Audiometer, model 1707 which was calibrated with Amplivox headphones according to ISO standards. Furthermore, an earplug was connected to this audiometer. The air-cylinder contained a reduction valve that regulated the air-flow to the electronic control system. The air-puff delivered through a tube of 2.5 mm in diameter lasted .5 sec. The amount of air delivered in one trial could be varied, but had a maximum of 1/2 liter. A more detailed description can be found in Lancioni and Hoogland (1980). Figure 1 shows a pictorial representation of this equipment.

Furthermore, an inflated red balloon (containing 3 1/2 liters of air) and a little soft brush were used for providing visual and tactile stimulation.

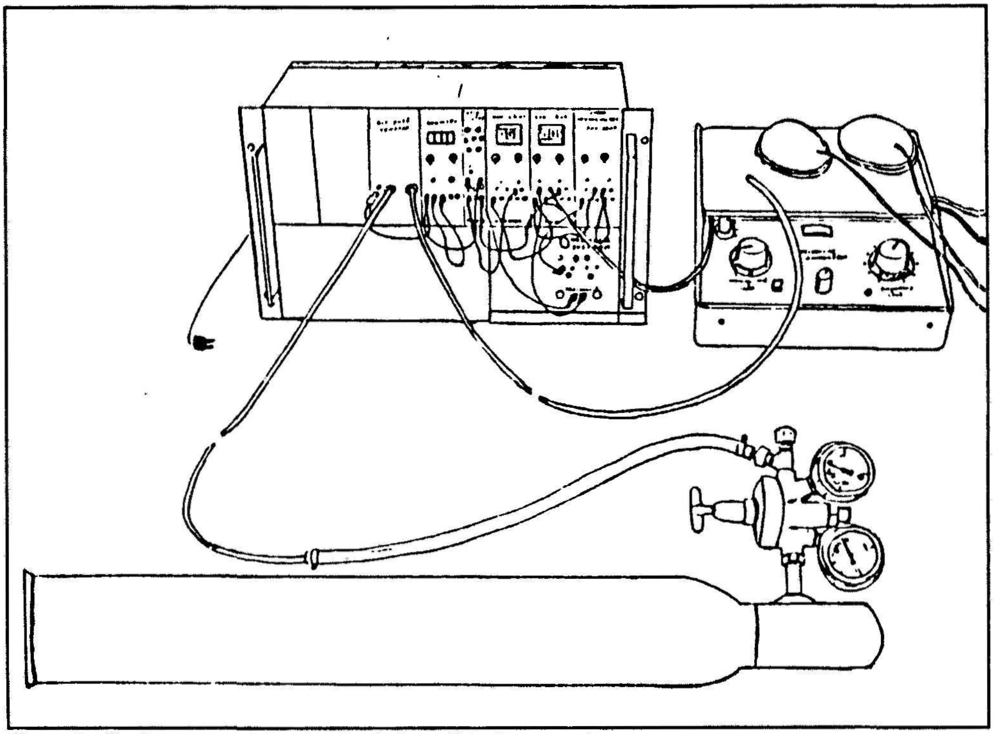

Figure 1. Equipment used for the classical conditioning procedures.

4.2.3.2. ASSESSMENTS AT 6 MONTHS

When the children were 6 months old additional learning rate measures were taken through two operant conditioning procedures. These relied on auditory and visual discriminative stimuli (S^Ds). The learning rate was defined as the number of trials required to achieve the criterion (see above).

The apparatus consisted of an electronic control device, connected to a) a loudspeaker or a screen (as auditory and visual discriminative stimuli), b) a box containing animated toys (reinforcement), c) a remote control box, through which the experimenter started trials and delivered reinforcement, and d) a cuing device informing the experimenter about the length of the trial (i.e., the period of time within which a response could be reinforced). Figure 2 shows a pictorial representation of the equipment used for the visual S^D condition. The equipment for the auditory S^D condition was essentially the same. The auditory S^D was produced via a loudspeaker, mounted on a stand (not shown in Figure 2). It consisted of two tones of 500 and 1500 Hz, alternating with a frequency of 8 per sec. The intensity could vary between 40 and 80 dB, but was normally

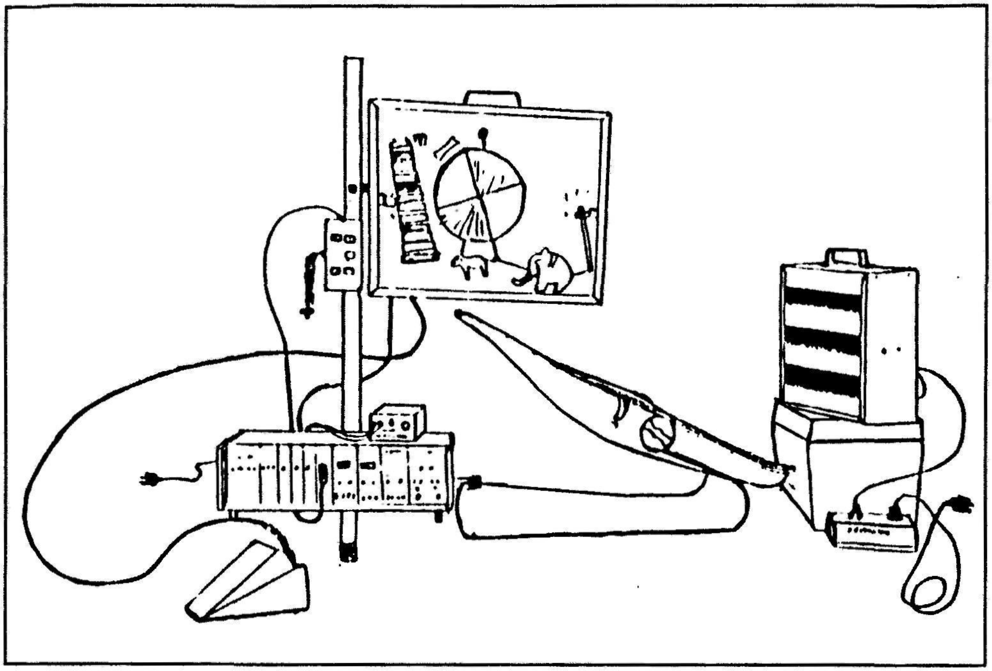

Figure 2. Equipment used for the 6 months' operant conditioning procedures with the visual S^D.

about 70 dB. The visual S^D was presented via a screen (measuring 40 by 40 cm). The screen was illuminated by two light bulbs and showed a pattern of black and white horizontal stripes. The reinforcement box was made of transparent plexi-glass. It could be turned around a horizontal spindle by operating a footpedal, connected to the box with a flexible steel wire. The turning of the box served to preclude intertrial reinforcement opportunities related to the intertrial responses of the subjects. When in operation the box showed animated toys, flickering lights and presented melodic sounds. In addition to this reinforcement box, several animated toys were available for reinforcement. They were applied according to identical criteria and conditions, and alternated with the aforementioned box.

4.2.3.3. ASSESSMENTS AT 18 MONTHS

Two measures were taken when the infants were 18 months old. That is, one learning rate measure obtained by an operant conditioning procedure, and a measure of the general mental and motor development provided by the Bayley Scales.

For the operant conditioning procedure an apparatus was employed, to which a remote control box and a cassette-recorder were connected. This apparatus provided discriminative stimuli (S^Ds), reinforcement and manipulanda for responding (see Figure 3 for a pictographic representation of the apparatus). It consisted of a rectangular box (with width, depth, and height of 80, 80, and 10 cm, respectively). On top of the back half of this box a housing (max height 50 cm) was positioned. This housing contained the reinforcing stimuli. The front of the housing was made of curved transparent

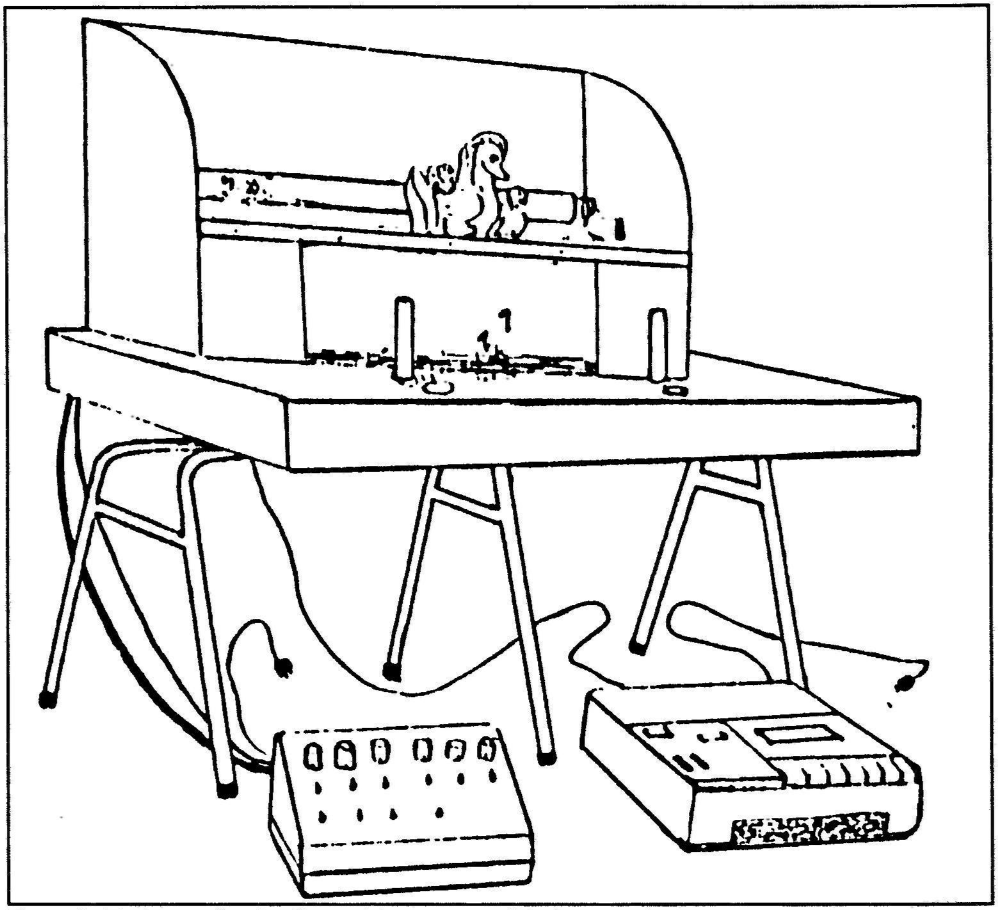

Figure 3. Equipment for the 18 months' operant conditioning procedure.

plexiglass. It contained colored sound-making animated toys, a loudspeaker for presenting children's songs, and a horizontally positioned rotating cylinder which presented colored pictures. When not activated the cylinder was invisible, while the toys and loudspeaker were still and silent. At the front-half of the top of the aforementioned box there were two round windows (diameter of 6.5 cm). They were positioned 30 cm apart (close to the manipulanda). Under these windows there were colored lamps, comprising the discriminative stimuli. The response manipulanda (push-pull-levers) were mounted 5 cm behind each window. For a correct response a minimal effort was needed with a displacement of at least 30 degrees. The remote control box and several switches for the control of the presentation of a) a discriminative stimulus by the activation of a red-colored lamp and b) reinforcement by the activation of the loudspeaker (music), cylinder (rotation), and animated toys.

In this study a Dutch translation of the Bayley Scales of Infant Development (BSID) was used, that is, the Bayley Ontwikkelingsschalen (BOS) (Meulen & Smrkovsky, 1983). These scales also provide norms adapted to the Dutch population. Both Bayley Scales, the Mental Development Index (MDI) (which contains items concerning perceptual and cognitive abilities), and the Psychomotor Index (PDI) (containing items measuring gross and fine motor development) were administered, according to the procedure described in the manual.

4.2.3.4. HOME OBSERVATIONS

Home observations were performed at the age of 6 months. The behavior of the children (responses) and environmental stimuli were recorded. The behaviors of the infants were the starting points of recording. In general terms, first the behaviors of the infants (responses), and secondly the stimuli at which the behaviors of these infants were directed, were recorded. The observations took place in the natural situation, with the restriction that the infant had to be awake and alert. The choice of the occasion and situation was left to the caretaker. The setting was mostly (a playpen in) the living-room. Generally, the period immediately after a feeding was chosen. The caretaker was instructed to act as he/she normally would in that situation. In addition, arrangements were made with the caretaker(s) that the circumstances during the observation were normal. That is, any time the observer or caretaker felt that the child or environment were not normal (e.g., the child was fussy; a visitor was present) the observation was repeated (the next day) or, with some delay, restarted.

The data were collected by means of a partial-interval recording system (Sulzer-Azaroff & Mayer, 1977). For the recording non-coded score sheets and timers connected to one or two earplugs were used. These timers produced pips to the earplug(s) to signal to the observers the beginning of a 5 sec interval. Each observation interval was followed by a recording interval. For the purpose of reliability one timer was connected to two earplugs such that the pips would always be synchronous. The score sheets were divided into two sections. One for scoring the behavior of the subjects, the other for environmental stimuli (for a sample of a score sheet, see Appendix A). The behaviors and stimuli are outlined in Table 3 and were scored in coded form. These behaviors and stimuli were taken from reports on observational child studies and were selected after extensive pilot observations carried out by the author in conditions comparable to those of the experimental group. Only those behaviors and stimuli were included that were judged relevant with regard to the aim of the study. That is, only those stimuli were included to which the child was attending or which had an observable impact on the behavior of the child. In addition, only those behaviors were included that were related to or could be related to environmental stimuli.

Table 3
Behaviors and Stimuli Recorded in the Home Observations

BEHAVIORS[a]	
	Looking, orienting and reaching
	Grasping (touching/holding) and manipulating
	Hand/finger(s) and or object in mouth
	Vocalizing (any oral-vocal sound other than laughing or crying)
	Smiling or laughing
	Fussing
	Crying
	Excitement (massive movement of arms and legs, whole body)
	Gross movement/displacement (rolling, creeping, pulling up to a standing position)
STIMULI[b]	
	Auditory (sound making, voice)
	Visual (toy, person)
	Animate (person)
	Inanimate (toy)
	Tactile-kinesthetic (picking up, swinging)

a) Within one observation interval more than one behavior could occur.
b) Stimulation could have one or a combination of the listed features, e.g., a toy-doll: visual and inanimate

On the basis of the aforementioned infant behaviors and environmental stimuli 12 observation items, comprising the raw material for statistical analysis, were composed. These items consisted of concurrent stimulus-response combinations or stimulus-response-classes (Marton, Minde, & Ogilvie, 1981), for example, 'Attention inanimate' refers to the looking/orienting and/or reaching at an object (play-material). In Table 4 the applied items are presented.

Table 4
Observation Items Composed of Behaviors and Environmental Stimuli as Presented in Table 3

Items	Behaviors and Stimuli
Attention inanimate	Looking, orienting and/or reaching at inanimate, auditory and/or visual (e.g., toy)
Attention animate	Looking, orienting and/or reaching at person
Mouthing	Hand/finger(s) and/or object in mouth
Vocalization	Any oral/vocal sound except laughing, crying, or fussing
Manipulation	Grasping, touching or manipulating objects by hand(s) or foot(feet)
General Excitement	Excitement, not directed at person
Excitement animate	Excitement directed at person
Smiling	Smiling or laughing
Movement	Gross movement, directed at person or object(s), or not directed at object or person (general)
Tactile-Kinesthetic	Person picking up or holding the infant
Following a request	Imitation of person or following simple request (e.g., 'roll-over')
Crying-fussing	Crying or fussing

4.2.4. Experimenter and procedures

4.2.4.1. ASSESSMENTS AT 4 MONTHS

A male experimenter (the author) carried out the conditioning procedures.
Conditioning took place when the subjects were awake, alert and quiet (state 3 or 4 according to Prechtl & Beintema, 1977), generally while the infants were drinking from their bottle and resting in their mothers' arms. On some occasions, however, the infants were lying in a baby chair. Experimental sessions were scheduled throughout the day.

The maximum number of conditioning (training) trials presented within one session was 15. The intertrial interval (ITI) ranged between 10 and 30 sec ($\underline{M}$ = 20 sec). Sessions lasted mostly 5 - 7 min, while sessions, interrupted because of state problems of the subject, could last 10 min.

Three classical conditioning strategies were used, each including baseline, conditioning, and testing trials. The three strategies varied with regard to the conditioned stimuli (CSs) that were used, that is, auditory, visual and tactile. These variations were applied on separate but consecutive days, in the aforementioned order. The unconditioned stimulus (UCS), an air-puff, was identical for all three variations. This air-puff had a duration of 1 sec, while its intensity could be varied, resulting in a maximum puff of 1/2 liter. It was delivered to the side of the infant's face (towards the corner of the eye) with the extremity of the tube held at a distance of about 10-15 cm.

4.2.4.1.1. Auditory conditioning procedure

A few trials (2-3) were presented in which the auditory stimulus, a tone of 2000 Hz with an intensity of 80 dB (SPL) and a duration of 1.5 sec, was applied. The aim was to ensure that the stimuli were not followed by any (or any recurrent) defensive reactions (e.g., blinking). They were presented through one headphone held close to the infant's ear. The headphone set was not employed to avoid irritation of the baby. After these initial trials a second set of trials was presented, during which the headphone was removed from the infants' ear. The experimenter delivered 2-4 airpuffs so as to regulate their intensity. This was done in order to obtain a criterion unconditioned response (UCR), that is, total closure of the eyes or repeated blinking, possibly accompanied by turning of the head. During conditioning, readjustments of the strength of the puff were possible so as to maintain a criterion response. Immediately thereafter conditioning started. During conditioning and testing the headphone and the tube to deliver the air-puff were held in the same position as during the baseline trials. The tone and air-puff were always presented on the same (normally the left) side of the subject.

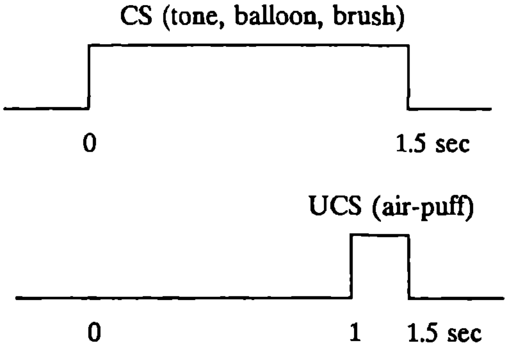

Figure 4. A training trial consisting of a 1.5 sec tone and a .5 sec air-puff. The last .5 sec of the tone overlapped entirely with the .5 sec air-puff.

Training trials were conducted according to a delayed paradigm. That is, a .5 sec air-puff (unconditioned stimulus) overlapped entirely with the final portion of the 1.5 sec tone (conditioned stimulus) (see Figure 4).

A maximum of 90 training trials could be provided. When the subjects showed two anticipatory responses, that is, complete and prolonged closure of the eye (a criterion response) immediately after the onset of the tone and before the onset of the air-puff, on two consecutive trials, testing trials were introduced. (Test and training trials were presented in ratios of 1 : 1 to 1 : 2.). A testing trial was a trial in which only the tone was presented, that is, the unconditioned stimulus was omitted. If after 25 trials no anticipatory responses had occurred one test trial to five training trials were momentarily introduced. They were planned to ensure against unnecessary prolongation of prompting, due to delayed responding of the subject. If during such test trials the subject showed conditioned responses (i.e., a criterion response within a test trial), the ratio of testing and training was changed into 1 : 1 to 1 : 2. The aforementioned ratios of 1 : 1 to 1 : 2 could momentarily be changed in 1 : 3 to 1 : 5 if the subjects failed to respond on the subsequent test trial(s). Training and testing continued until a learning criterion had been reached, or after 90 training trials if the subject failed to reach the learning criterion. This learning criterion was defined as 4 conditioned responses on as many consecutive test trials or 5 conditioned responses on 6 consecutive test trials, without responding between trials (Lancioni et al., 1989). The number of training trials the subjects needed to reach criterion constituted the learning rate. One day after the termination of this strategy (when the conditioning criterion was reached or a maximum of 90 training trials had been applied) the next procedure was implemented.

4.2.4.1.2. Visual conditioning procedure

A few trials (2-3), applying the visual stimulus, an inflated red balloon, were presented for 1.5 sec, with the same aim as in the previous strategy. Before and in between presentations the balloon was held (by the experimenter) behind the head of the subject (invisible to the subject). Then the balloon was moved over the head and brought in front of the eyes, at a distance of about 20 cm. The movement was carried out slowly so that it would facilitate orienting without provoking defensive reactions.

During conditioning and testing the balloon, which constituted the conditioned stimulus (CS), was used as in baseline. During intertrial intervals it was held behind the head of the subject. Training trials were conducted according to a delayed paradigm (i.e., following the same rules as in the Auditory Conditioning procedure, see also Figure 4). For the presentation of the balloon the experimenter was cued by an auditory signal, presented through an earplug, 2.5 sec before the presentation of the unconditioned stimulus (i.e., the airpuff). All the other conditions for training and testing were identical to those described in the previous strategy, except that the maximum number of training trials available was 75. One day after the termination of this strategy the next strategy was implemented.

4.2.4.1.3. Tactile conditioning procedure

As for the previous conditioning procedure, the tactile stimulus was initially presented 2-3 times with the purpose of ensuring that it alone would not create defensive responses. The tactile stimulus consisted of a light strike with a brush on the upper left side of the subject's head. The strike lasted about 1.5 sec, and covered an area of the skull of about 5 cm.

Immediately after the baseline trials training trials were presented. During conditioning and testing the tactile stimulus (CS) was used as in baseline. Training was executed according to a delayed paradigm, identical to that employed within the previous conditioning procedures. For the presentation of the strike the experimenter was cued as in visual conditioning. Yet, the cue started 1 sec before the presentation of the unconditioned stimulus (i.e., the airpuff). All training and testing conditions were as in the previous procedures, except that the maximum number of training trials available was 45.

4.2.4.2. ASSESSMENTS AT 6 MONTHS

A male experimenter (the author), an observer and the caretaker of the subject were employed in the execution of the conditioning procedures.

Conditioning took place when the subjects were alert and quiet, generally sitting/lying in a babychair. During an experimental session the position of the experimenter was behind the baby-chair and out of the visual range of the subject. He held a control box through which he determined the presentation of the stimuli. The caretaker was sitting next to the chair of the baby (on the right side of the baby). His/her face was directed towards the infant. The observer, with a score-sheet and a stopwatch, was sitting opposite the experimenter and facing the infant. Experimental sessions were scheduled throughout the day (a maximum of four per day), with intervals of at least half an hour between them. The maximum number of training trials presented within one session was 10. The intertrial intervals were variable and controlled by the experimenter. They depended on the state and responses of the infant. Mostly the intervals were between 10 and 30 sec. The mean length of these intervals was 21 sec. A session lasted generally between 6 and 10 min.

Two operant conditioning strategies were used, each including baseline, conditioning (trials with prompts), and probing trials (trials without prompts). The two strategies varied with regard to the discriminative stimuli (S^Ds) used, that is, visual or auditory. They were implemented on separate but consecutive days, first the visual and then the auditory procedure. The reinforcing stimuli were identical for both procedures (see description of materials). They were presented for 3 sec at each instance.

4.2.4.2.1. Visual discriminative stimulus procedure

During baseline the position of the subject, caretaker, experimenter and observer were as described. The assessment room was dim-lit. The box for the presentation of the discriminative stimulus (S^D) was positioned in the same horizontal plane as the subject's head, just behind the baby-chair at a distance of 50 cm from the subject's face (see Figure 2 for a pictorial representation of material).

In the baseline trials (2-3) only the discriminative stimulus was presented (for 6 sec). Each trial was started when the baby was looking at the caretaker (the subject's head was turned to the right). The caretaker attempted to maintain the subject's attention by looking at and softly speaking to the infant. These trials were planned to ensure that the infant would orient to the discriminative stimulus and at the same time would adjust to the experimental situation. Immediately thereafter the first series of 10 training trials were started. During these trials the positions of the subject, caretaker, experimenter, observer and materials were as in baseline. The response to be trained was head orientation toward the reinforcement box (an upward headturn), within 6 sec after the onset of the S^D. Intertrial responding reset the intertrial interval. Training was carried out according to a one discriminative stimulus, discrete trial operant conditioning paradigm (Lancioni, 1980; Reese & Lipsitt, 1970). Initially, responding was facilitated with the help of physical prompting (directed at bringing about the target response) and reinforcement. During a trial (6-sec presentation of the S^D), the caretaker helped the subject to turn towards the reinforcement box by touching the upper-side of the baby's head or by executing some pressure against the baby's head with his/her hand. Immediately after a headturning response reinforcement was administered (3 sec). That is, either the reinforcement box turned towards the infant and activated, or one of the animated toys was brought above the infant's head and activated. Immediate prompting (within 1 sec after onset of S^D) was provided in the first three trials of the first session and in the first trial of the following session(s). From the fourth trial in the first session to the sixth trial of the second session prompting was gradually delayed 4-5 sec (the experimenter was cued by an auditory signal through an earplug). Thereafter, probing trials were normally interspersed. These probing trials (for which no prompts were available) were presented at an average of one every three trials. This ratio was changed to 1 : 1 if two discriminative responses (headturning responses during training trials after the onset of the S^D and before the provision of a prompt) during two consecutive training trials or one headturning response during a probing trial occurred. If on two consecutive probing trials headturning responses occurred, only probing trials were presented. The learning criterion was achieved when the subjects had shown seven positive responses over as many or a maximum of nine consecutive probing trials. However, a maximum of 60 trials could be provided. One day after the termination of this procedure the procedure with the auditory discriminative stimulus started.

4.2.4.2.2. Auditory discriminative stimulus procedure

During baseline trials the positions of subject, caretaker, experimenter and observer were as in the previous procedure. The loudspeaker for the presentation of the auditory S^D (6 sec) was installed above the infant's head, at a distance of 30 cm. The reinforcement box was positioned in the same plane as the subject's head, at a distance of about 50 cm, on the left side of the baby-chair.

Baseline trials were carried out as in the visual S^D procedure. Baseline was followed by the first series of training trials. During training as well as during probing, conditions were the same as for the visual S^D procedure.

4.2.4.2.3. Home observations

At the age of 6 months also home observations were carried out. The observations were carried out by one observer, incidentally by two (the second acted as reliability observer). These reliability observations were randomly spread over children and ages and covered, timewise, more than 11% of the total study period. The observers had trained themselves, before the study started, with different subjects (comparable to the experimental subjects) and in similar conditions. Each home observation consisted of two sessions of 20 min each, separated by a little break (5-10 min). In case of state problems and the like (see section 4.2.3.4) additional breaks were applied.

4.2.4.3. ASSESSMENTS AT 18 MONTHS

One male and two female experimenters and the caretaker were employed to administer the Bayley Ontwikkelingsschalen and the operant conditioning procedure.

Conditioning took place with the subjects standing or sitting in front of the apparatus. They were always within reach of the two response-levers. The caretaker was positioned immediately behind the subject. He/she was instructed not to interfere with the procedure, except when explicitly asked by the experimenter. His/her task was to perform the introduction (demonstration) trials, to provide the prompts and to withdraw the hand(s) of the subject from the response-lever(s) during intertrial intervals, whenever needed. The experimenter, while operating a control panel, and the observer were behind the subject and caretaker. Experimental sessions were scheduled throughout the day, a maximum of three per day, with intervals of at least one hour between them. The maximum number of training trials presented within one session was 10. The intertrial intervals were variable and were controlled by the experimenter. Most intervals were between 10 and 30 sec, while the average length was about 20 sec. A session lasted between 6 and 10 min. The strategy included introduction, conditioning (training) and probing trials.

During the introduction the positions of the subject, caretaker, experimenter and observer were as described above. A few (3-5) demonstration trials were presented, while the caretaker performed the required response (the infant was watching, but prevented from responding). That is, a discriminative stimulus was presented for 5 sec (either left or right window was illuminated for 5 sec) or less if a response occurred. The caretaker had to push or pull the corresponding lever (the lever just behind the window) after onset and before the offset of the S^D and reinforcement followed, thus had to make the correct choice out of two, representing a discriminative response (see Figure 3 for a pictorial representation of material).

After the aforementioned introduction trials training was applied according to a two-discriminative stimuli, discrete-trial operant conditioning paradigm (Lancioni, 1980). The maximum number of training trials was 60. Training trials consisted of the presentation of the S^D (5 sec), during which the subject was prompted (by physical guidance and/or verbally) to perform the required response before the offset of the S^D. These training trials were interspersed with probing trials, that is, trials without prompts. Performance of the required response both within a training and probing trial was reinforced. Reinforcement consisted of activation of the reinforcing apparatus, combined with praise, for 5 sec. Incorrect responses (i.e., manipulation of the lever not corresponding with the lighted window) were always followed by a correction trial. Correction trials were identical to training trials with physical prompts. Training and fading out of the prompts occurred according to the following guidelines. The first four trials involved immediate prompting. The fifth trial was a probing trial. Thereafter the ratio of one probing trial and five training trials continued, but the prompt was delayed for 3 sec with regard to the onset of the S^D. If during a probing trial a correct response occurred, the aforementioned ratio (of 1 : 5) was changed to 1 : 2. The change was also made if a correct response was made after presentation of the S^D and before a delayed prompt. After two correct responses in two consecutive probing trials, or two correct responses during two consecutive training trials, only probing trials were applied. However, after two incorrect responses in two consecutive probing trials the sequence of probing and training trials (with a ratio of probing : training = 1 : 2) was reinstated. The learning criterion consisted of five correct responses in five consecutive probing trials or six correct responses out of seven consecutive probing trials.

4.3. Results

4.3.1. Outline and data preparation

Data are presented in the following sections for the total group of 48 infants. Data collection took place from the age of 4 to 18 months. Analyses are also presented for the two separate subgroups, that is, the high risk and handicapped (n=36) and the normal infants (n=12). General characteristics of the infants are described in Section 4.2. (Method).

As mentioned before, the infants were tested at the ages of 4, 6 and 18 months. The measures taken at the ages of 4 and 6 months (predictor measures) were used to predict the measures taken at the age of 18 months (outcome measures).
The following variables were involved: Classical Auditory (scores on classical conditioning task with auditory CS), Classical Visual (scores on classical conditioning task with visual CS), and Classical Tactile (scores on classical conditioning task with tactile CS); Operant-6 Visual and Operant-6 Auditory, that is, scores on the six months' operant conditioning tasks, with visual and auditory discriminative stimuli, respectively, and eight home observation items. At the age of 18 months the following variables were obtained: Operant-18 (operant conditioning), and finally the BayleyMS and BayleyMR (scores on the Mental and the Motor Scales of the Bayley Scales of Infant Development, respectively).

In a number of infants only part of the investigations could be performed, and thus for some of them data were incomplete (see also Tables 5 and 6). That is, for three high-risk infants none of the three operant conditioning procedures was performed because of equipment failure. For three other high-risk infants the entire Bayley assessments (in one subject) or parts of the Bayley assessments were not carried out due to severe retardation or motor handicaps. Three normal infants did not complete all the assessments because of family reasons such as change of residence. Of these three subjects one did not undergo the 18-months' assessments. For the other two subjects no Bayley assessments were executed. As a result of the missing data, different analyses were performed on variable numbers of subjects. For each analysis the number of subjects entered will be indicated.

In addition to the missing data, a number of instances occurred in which subjects underwent conditioning but did not reach the conditioning criterion within the allowed number of trials (as defined before the beginning of the experiments; see Section 4.2.). These subjects, which did not succeed conditioning in one or more procedures, will be referred to as "unsuccessful subjects" or "conditioning failures" (see Wachs & Smitherman, 1985). In all these cases scores were allotted to make data analyses feasible. A rationale for inclusion of those conditioning failures and the criteria followed for attribution of their scores are provided below.

Table 5
Subjects Undergoing and Succeeding* the Different Assessment Procedures. For Each Learning Procedure the Maximum Allowed Number of Training Trials is Shown

	Max no trials	High-risks Undergoing	High-risks Succeeding	Normals Undergoing	Normals Succeeding
Classical Auditory	90	36	36	12	12
Classical Visual	75	36	35	12	12
Classical Tactile	45	36	36	12	12
Operant-6 Visual	60	33	24	12	11
Operant-6 Auditory	60	33	30	12	12
Home observation		35		11	
Operant-18	60	33	28	11	8
BayleyMS		35		10	
BayleyMR		33		10	

Note. Variables are described in the text
* Reaching the conditioning criterion within the maximum allowed number of trials.

In most instances, the data used for analyses were the raw scores on the measures applied in this study. That is, these raw scores were for the learning assessments the number of training trials a subject needed to reach a learning criterion (learning rate) and for the home-observations the raw scores on a number of items (see method section). When a subject (for the classical and operant conditioning procedures) did not reach the criterion within the maximum number of trials provided, data adaptation procedures were applied as outlined in the following paragraph. With regard to the Bayley assessments the raw scores on the Mental and Motor Scales were chosen instead of developmental quotients. Justification is found in the following considerations: 1) all the children were assessed at the same age, 2) emphasis in this study was on gathering predictive and not normative data, and 3) avoidance of loss of variance due to transformation into normalized scores (Vietze, McCarthy, McQuiston, MacTurk, & Yarrow, 1983).

As is described in Section 4.2., a limited number of training trials was provided for each learning procedure. This number was set before the experiment started and was determined after pilot investigations on normal infants had taken place. In other words, if a subject did not reach the conditioning criterion within the maximum number of training trials allowed, training and testing on the same procedure was stopped, assuming that the failure to reach criterion would mean that the subject was not able to learn under the available conditions. Moreover, it was considered that not reaching a learning criterion in the available trials would indicate a low learning ability. Furthermore, the rationale for including the unsuccessful subjects was based on arguments derived from Bathurst and Gottfried (1987) and Wachs and Smitherman (1985). In short, these authors maintain, that not completing a testing procedure is not a matter of coincidence, but can be ascribed to subject characteristics, that is, resulting in this experiment in slow learning and thus in developmental delays. In the present study 'unsuccessful' may mean successful in all but one conditioning procedure, or in other words could be defined as partially successful. In fact, no subject was totally unsuccessful, that is, did not reach the conditioning criterion within the available training trials in any of the presented classical and 6-months' operant conditioning procedures. Therefore it was decided to include also the subjects who were not successful in learning (or conditioning failures), and to give them scores according to the following procedure. The choice of the maximum number of trials used did not seem appropriate (in fact, they did not represent a learning rate). The magnitude of the substitute score was then, rather arbitrarily, chosen to be 5 to 14 units greater than the maximum number of trials allowed. For each subject the score was allotted through the use of a table of random numbers, taken from Blalock (1979). It was not chosen to allot identical scores for each unsuccessful subject, because this would result in an asymmetric score-distribution. In addition, the range of these allotted scores is rather limited (compared to the expected 'real' scores). This was done to approach normally distributed scores. Moreover, the followed procedure very probably would provide for a conservative estimate of the real scores to be expected. The number of subjects who underwent and succeeded conditioning are shown in Table 5. The raw scores, including the allotted scores, of all subjects on the main experimental variables are shown in Table 6. The table shows that in total for the classical procedures on only one occasion a score was attributed; for the Operant-6 Visual the number was ten, and the Operant-6 Auditory three. The Operant-18 appeared to be very sensitive to conditioning failure, that is, eleven subjects did not condition.

Summarizing the data preparations for statistical analyses, complete data on 39 and incomplete data on 9 subjects were available. Thus, the data of the latter subjects were entered only in some of the statistical analyses.

At first, descriptive statistics are reported for the two subgroups as well as for all the subjects together (total group of subjects). The data are presented first for the different predictor measures and thereafter for the outcome measures so as to compare the scores with those found in other reports. Moreover, results of statistical analyses are presented to describe relationships among the variables, the main goal of investigation.

In addition, an attempt was made to compare the successful with the unsuccessful high-risk subjects quantitatively as well as qualitatively. This was done to get more insight into the question why some subjects were not able to learn within the maximum number of available trials. Furthermore, all the subjects were classified using the scores on the Bayley Mental Scale to provide possible evidence on the predictive and concurrent validity of the used learning measures and to investigate whether some selected predictors could be used for classification of individual subjects.

Table 6
Raw Scores of the Subgroups of High-risk and Normal Subjects on the Main Experimental Variables. An Asterisk Indicates an Allotted Score. A Score Is Not Reported When It Was Not Available due to Equipment Failure, and Change of Residence, Severe Retardation or Impairment of the Subject (missing data)

	4 months			6 months		18 months		
	Classical			Operant-6		Operant18	Bayley	
Nr	Aud	Vis	Tac	Vis	Aud		MS	MR
				HIGH-RISK GROUP				
1	17	13	5				123	51
2	13	36	7				134	52
3	22	21	8				120	54
4	23	11	5	12	13	20	143	59
5	22	9	5	12	5	22	147	60
6	35	38	5	21	11	44	127	52
7	20	22	9	13	32	20	121	51
8	18	34	4	18	38	19	127	50
9	22	18	12	74*	12	13	121	58
10	22	12	5	68*	27	65*	120	55
11	21	33	4	12	22	41	116	51
12	17	32	6	9	23	13	136	49
13	18	14	4	22	7	35	126	55
14	40	19	4	28	3	71*	113	44
15	12	12	5	11	37	65*	122	
16	10	12	5	67*	21	74*		
17	13	10	4	20	45	33	121	52
18	37	18	8	66*	71*	72*	75	32
19	41	22	8	17	19	32	130	
20	12	15	12	8	5	23	98	44
21	18	12	5	65*	49	66*	115	51
22	18	18	7	30	22	33	123	50
23	14	17	4	35	9	45	123	53
24	18	13	8	8	6	4	121	52
25	22	12	5	28	41	51	116	51
26	25	20	7	11	54	26	111	51
27	20	34	11	70*	34	18	119	53
28	35	37	22	72*	42	72*	85	28
29	25	19	7	12	14	15	119	48
30	35	20	12	30	13	44	113	48
31	25	25	12	70*	30	24	100	33
32	49	87*	8	55	67*	24	111	44
33	18	19	5	7	11	7	130	53
34	18	20	5	13	22	16	129	50
35	28	23	13	71*	42	68*	103	44
36	17	10	7	12	67*	48	104	44

Continuation table 6

	4 months			6 months		18 months		
	Classical			Operant-6		Operant18	Bayley	
Nr	Aud	Vis	Tac	Vis	Aud		MS	MR
				NORMAL GROUP				
37	19	14	7	41	47	22	130	45
38	15	11	3	16	13	22	122	51
39	20	20	5	16	18	17	135	52
40	23	22	5	31	6			
41	9	5	4	74*	22	12	128	54
42	25	25	18	23	31	29	123	51
43	24	22	9	22	29	69*		
44	49	27	8	13	29	68*		
45	6	25	4	13	31	65*	124	53
46	23	18	5	21	20	35	121	53
47	19	14	5	16	31	7	135	53
48	22	10	5	18	8	27	121	53

Note. Aud = Auditory; Vis = Visual; Tac = Tactile.

4.3.2. Description of the predictor measures

In Table 7, the mean scores and ranges on the predictor measures of the two subgroups are summarized. None of the variables showed statistically significant differences between the two subgroups applying t-tests[1], except the two observation variables Vocalization and Crying-fussing (p < .05). The scores show quite some variability between infants, that is, some infants learned the association in less than 10 trials, while some others failed to learn in 90 trials. These findings cannot easily be compared with those of other investigations, because of the different aims of the studies and procedures and subjects used. In two studies, applying procedure, stimuli, and criteria identical to those applied in the auditory classical conditioning procedure of the present study (Lancioni & Hoogland, 1980; Lancioni et al, 1985), highly similar scores were found. Concerning the operant procedures no studies were found where similar procedures and subjects were used aiming at a similar goal, that is, to establish learning rate scores.

1 Mann-Whitney U tests (performed on rank ordered scores) yielded identical results.

Table 7
Descriptive Statistics: Predictor Variables

	Subjects							
	High-risks				Normals			
Variable	Mean	S^D	Range	n	Mean	S^D	Range	n
Classical Auditory	22.8	9.2	10-49	(36)	21.2	10.6	6-49	(12)
Classical Visual	21.9	14.0	9-87	(36)	17.7	6.9	5-27	(12)
Classical Tactile	7.3	3.7	4-22	(36)	6.5	4.0	3-18	(12)
Operant-6 Visual	32.3	24.9	7-74	(33)	25.3	17.3	13-74	(12)
Operant-6 Auditory	27.7	19.2	3-71	(33)	23.7	11.6	6-47	(12)
Attention inanimate	26.6	15.9	2-67	(34)	25.3	14.4	8-48	(11)
Attention animate	41.9	16.2	15-82	(34)	38.0	18.8	18-76	(11)
Mouthing	19.4	14.8	0-53	(34)	14.6	8.7	2-29	(11)
Vocalization	8.9	11.7	0-59	(34)	3.7	4.3	0-12	(11)
Manipulation	36.1	17.1	10-74	(34)	39.6	24.8	1-72	(11)
Smiling-excitement	6.2	5.6	0-18	(34)	6.4	4.4	1-15	(11)
Gross movements	3.6	6.4	0-28	(34)	3.6	6.2	0-16	(11)
Crying-fussing	8.1	12.2	0-56	(34)	2.8	3.2	0- 9	(11)

Note. Figures in parentheses indicate the numbers of subjects undergoing the procedure.

Furthermore, it can be noticed that within each subgroup and within both conditioning procedures (classical and operant) there is a decrease in mean scores according to the order of presentation, due to generalization learning, stimulus characteristics, or both.

Home observations were performed by two observers. Because some items turned out to have a very low frequency of occurrence, they were either dropped (i.e., Following response) or a composite score was computed by combining some items, that is, Smiling and Excitement animate (see also Barrera, Rosenbaum, & Cunningham, 1986; Wachs, 1984). As a result the scores on eight (new) variables were computed (see Tables 3, 4, 7 and 8). These were used to supplement the learning measures in predicting outcome measures. A description of the observation items was provided in the method section (subsection measures and material).

With regard to reliability, an agreement was recorded when both observers had scored the occurrence or the non-occurrence of a behavior within the same observation interval (duration of 5 sec). If within an interval only one observer had scored a behavior this was considered a disagreement. Within the same interval more than one behavior could occur (Sulzer-Azaroff & Mayer, 1977). The number of behaviors within an interval and across subjects varied between 1 and 5 with a mean of 1.8. The reliability sessions covered more than 11% of the total observation time and were spread over the entire study period. Several estimates of interobserver agreement were computed, because of the variability in occurrence among the behaviors, tha is, some behaviors were observed in most, while others in only some intervals (see also Table 7). The percentages of agreements between observers on the different behaviors during reliability sessions (outside the study sessions, but covering the study period) varied between 88 and 99% when all the intervals were included (total agreement). In addition, kappa coefficients were computed (Oud & Sattler, 1984), which ranged between .66 and .84. Furthermore, occurrence and non-occurrence agreement was computed (Hartmann, 1984; Sattler, 1988) (see Table 8). Occurrence agreement was computed over those intervals in which at least one of the two observers recorded the occurrence of a behavior. Non-occurrence agreement took into consideration only those intervals in which either one or both of the observers recorded the non-occurrence of a behavior. Because reliability checks took place over many intervals and some behaviors occurred in a relatively small proportion of those intervals, occurrence agreement as well as kappa represent conservative indexes of interobserver reliability. For most behaviors reliability could be considered modest to satisfactory, particularly when one considers the large number of observation categories. For the categories Gross movements and Crying-fussing (which occur very little) reliability is low.

It has to be noted that the mean frequency of occurrence of the behaviors was quite variable (see Table 7). Furthermore, the frequency of occurrence as tabulated in Table 7 shows that, proportionally, the infants were largely engaged in looking at animate stimuli and in manipulating objects (Manipulation). An animate stimulus was generally the primary caretaker (Attention animate). While manipulating objects, the infant would normally pay visual attention to the manipulated object. This was not scored as Looking at inanimate stimuli. To a lesser, but still substantial degree, they were paying attention to inanimate stimuli (objects) by looking, orienting or reaching (Attention inanimate). The frequency of the category Mouthing was almost as high. This behavior could occur in overlap with other behaviors (except Vocalization). The remaining categories were observed relatively less often. In Table 7 it can also be seen that the mean frequencies of most variables did not differ among the subgroups.

Table 8
Percentages of Occurrence, Non-occurrence and Total Agreements and the Kappa Coefficient Between Two Observers on the Different Behaviors

BEHAVIOR	Occurrence %	Non-occurrence %	Total %	Kappa
Attention inanimate	75	87	91	.79
Attention animate	74	81	88	.76
Mouthing	70	96	96	.78
Vocalization	74	94	94	.82
Manipulation	77	90	93	.84
Smiling-excitement	58	93	93	.70
Gross movements	50	99	99	.66
Crying-fussing	75	96	96	.82

4.3.3. Relations among predictors and differences between subgroups

Relations among predictor measures were computed by means of Pearson-r correlations. The correlations were computed pairwise, that is, only those subjects were entered into the analysis for the procedure on which scores were available on that pair of measures, including attributed scores. These analyses were performed to assess relationships among predictor variables in order to investigate possible overlap in content. An additional aim was to investigate possible differences between the two subgroups. Correlation matrices (including significance levels) among the predictor variables (except the home observation variables) for the different groups are presented in Table 9.[2]

For the total group the three classical measures are correlated significantly (correlations of .27 and greater, all $p < .05$, $\underline{n} = 48$). Also the two operant measures had a statistically significant correlation ($\underline{r} = .31$, $p < .05$, $\underline{n} = 45$). Most other correlations between conditioning variables (including the correlations among the classical and operant conditioning variables) are not significant and low, except for one, that is, the Classical Tactile measure. With regard to the home-observation variables, for this group only the variables Attention inanimate, Vocalization and Gross movements are significantly correlated (values of -.26, +.52, +.33, respectively) with the Classical Visual measure. In addition, Attention animate was correlated significantly ($\underline{r} = .32$) with the Classical Tactile measure. Several of the home-observation measures showed to be interrelated. The aforementioned correlations among home-observation variables are not presented in a table. The correlations in the high-risk group are similar to those in the total group, but a somewhat larger number is significant. In the normal group the relationships are generally modest to high, but not significant. No explanation can be provided for the large and significant negative correlation between both visual conditioning measures. Only four correlations between the home-observation measures and other predictors were statistically significant. Most correlations were low and did not point to meaningful relationships (not reported here).

Summarizing the results on the predictive variables presented thusfar, it could be concluded that proportionally more high-risk than normal children failed to reach the conditioning criterion (see Table 5). That is, only one of the children did have problems in achieving the conditioning criterion with one classical conditioning task. This finding, which is in agreement with earlier studies (Lancioni & Hoogland, 1980; Lancioni et al, 1985), is very important when one considers the generally reported high rate of subject loss in infant habituation and learning studies (Wachs & Smitherman, 1985). The operant tasks led to several failures in the high-risk group (11 failures in total, most of them occurring in the visual S^D procedure), which is consistent with the findings in habituation and learning studies, while only one failure was seen in the subgroup of normals. On the one hand, this makes the operant conditioning procedure less suitable for studying (high-risk) infants, because completion of a test (and not subject loss) is normally required. On the other hand, the fact that subjects fail to complete a procedure due to the inability to reach the conditioning criterion within a relatively high number of trials (compared to normal subjects) could be considered an indication of slow learning.

---

2 Spearman rho's (Pearson correlations on rank-ordered data) revealed similar results, with the same statistical significance, except that most correlations were somewhat lower.

This latter hypothesis is investigated later in this paper. For all the learning tasks the high-risk children needed more trials to achieve the criterion than the normals, although for none of the differences statistical significance is reached. The range of scores in the high-risk subgroup is also broader. Several measures, particularly those from some of the classical and operant conditioning procedures, were clearly related. This indicates that the classical procedures and to a lesser degree also the operant procedures seem to tend to similar learning scores across the subjects. Across the classical and operant procedures the scores seem to be less related. The latter low correlations between the classical and operant measures point to a potentially meaningful use in multiple regression analyses. The relation between home-observation variables and the other predictor-measures is, based on thusfar presented analyses, not meaningful.

Table 9
Pearson-r Correlations Among the Predictor Learning Measures for the Different Groups

Variable	Classical Visual	Classical Tactile	Operant-6 Visual	Operant-6 Auditory
		TOTAL GROUP		
Classical Auditory	.51 (48)	.34 (48)	.13 (45)	.18 (45)
Classical Visual		.27 (48)	.14 (45)	.31 (45)
Classical Tactile			.37 (45)	.19 (45)
Operant-6 Visual				.31 (45)
All correlations ≥ .27 are significant (p < .05)				
		HIGH-RISK GROUP		
Classical Auditory	.54 (36)	.32 (36)	.25 (33)	.21 (33)
Classical Visual		.24 (36)	.21 (33)	.31 (33)
Classical Tactile			.47 (33)	.14 (33)
Operant-6 Visual				.33 (33)
All correlations ≥ .31 are significant (p < .05)				
		NORMAL GROUP		
Classical Auditory	.50 (12)	.39 (12)	-.35 (12)	.04 (12)
Classical Visual		.49 (12)	-.57 (12)	.21(12)
Classical Tactile			-.09 (12)	.38 (12)
Operant-6 Visual				.09 (12)
None of the correlations are significant except the correlation of .57 (p < .05)				

Note. Figures in parentheses indicate the numbers of subjects entered into the analysis.

4.3.4. Description of the outcome measures

The following outcome measures, obtained through the assessments at 18 months of age, were involved in data analyses: operant learning rate (Operant-18) and Bayley Mental (BayleyMS) and Motor (BayleyMR) scores. The scores for the single procedures are reported for the two subgroups (infants undergoing a procedure), that is, the group of high-risk and the group of normal infants. The means and ranges for both subgroups (and differences between the subgroups and respective t-values) are presented in Table 10.[3] The Bayley scores were compared with those available from the standardization sample of the test in order to provide some developmental characteristics of the study samples used.

Table 10
Descriptive Statistics: Outcome Measures

Variables	High-risks				Normals				t
	Mean	S^D	Range	n	Mean	S^D	Range	n	
Operant-18	37.1	21.7	4- 74	(33)	33.9	22.8	7- 69	(11)	.40
BayleyMS	118.3	14.4	75-147	(33)	126.6	5.7	121-135	(9)	-2.66
BayleyMR	49.2	7.1	28- 60	(33)	51.7	2.7	45- 54	(9)	-1.64

Note. Figures in parentheses indicate the numbers of subjects entered into the analysis.

With regard to the operant conditioning procedure, it showed that the high-risk infants needed on average more trials to reach conditioning than the normal infants. The difference was not statistically significant. For both groups the scores were quite spread. The scores on both Bayley subscales are lower for the high-risk group than for the other subgroup. However, only the difference in group-means of the Mental subscale reaches statistical significance. However, the subjects of the normal group showed much less variability in scores. Quite a number of infants (more than 20%) did not reach the conditioning criterion for the operant conditioning procedure within the maximum available number of trials (see Table 5).

With regard to the Bayley scores of the subgroup of normal children (although very small), it could be maintained that these do not deviate from the scores of the population studied for the Dutch version of the Bayley (the Bayley Ontwikkelingsschalen, BOS 2-30, Meulen, B.F.van der, & Smrkovsky, M., 1983), the version actually used in this study. In fact, for the population of the 18-months-old children (forming the BOS 2-30 standardization sample), the raw score mean of both the mental and motor scales were 125.68 and 52.37, respectively. The group of high-risk mental as well as on the motor scale. These latter scores showed to be lower (reaching statistical significance) than the

3 Analyses with Mann-Whitney U statistics on the means of the rank-ordered scores (on the data used in Table 10) and Spearman rho correlations (on rank ordered scores of the data used in Table 11) showed the same results.

scores infants, however, showed somewhat lower scores, both on the of the standard-normal group participating in the BOS 2-30 study.[4]

With regard to the scores on the operant conditioning procedure no comparison could be made with other research findings, because of unavailability of comparable studies.

4.3.5. Relations among outcome measures

Relations among these measures were computed by means of Pearson-r correlations. These correlations were computed pairwise, that is, only those subjects were entered into the analysis for the procedure on which scores were available on that pair of measures, including allotted scores. The correlations were computed to find out how relationships in the experimental groups would compare with those found in literature. In addition, it was considered relevant to evaluate relationships among the different measures for information on concurrent validity.

Table 11
Correlations Among the Outcome Measures for the Different Groups

Variable	BayleyMS	BayleyMR
Total group		
Operant-18	-.54 (41)	-.42 (39)
BayleyMS		.82 (42)
High-risk group		
Operant-18	-.53 (32)	-.45 (30)
BayleyMS		.86 (33)
Normal group		
Operant-18	-.54 (9)	.09 (9)
BayleyMS	-.15 (9)	

Note For _n_ > 30 a correlation of 30 is significant (_p_ < 05) None of the correlations for _n_ = 9 are significant
In parentheses the number of subjects entered into the procedure is indicated

Inspection of Table 11 reveals that for the total group as well as for the high-risk group Pearson-r correlations between all outcome measures are statistically significant in the expected directions, and rather high.[3] That is, there is a positive correlation among the Bayley Mental and the Bayley Motor Scale. In addition, negative correlations between both the latter scales and the operant procedure are found. Moreover, the correlation between the Operant-18 and the BayleyMS is always higher than the correlation between the first measure and the BayleyMR.

4 The author thanks van der Meulen and Smrkovsky for performing t-tests for comparison of the Bayley data of the study presented here and their own normalization data In addition, they are acknowledged for providing comments on those comparisons

Summarizing these descriptive statistics, proportionally as many high-risk infants as normals failed to reach the conditioning criterion in the 18 months' operant conditioning task. That is, more than 20%, which could be considered a high percentage. On the average, the subjects of the high-risk subgroup needed more trials than the normal infants to achieve the criterion on the operant conditioning task. Overall, the high-risk infants achieved somewhat lower scores on both Bayley measures than the normal infants. The three outcome measures were clearly correlated.

4.3.6. Relations among predictor and outcome measures

The main goal of this study was to investigate the efficacy of some predictive models based on variables grouped into several sets. This goal will be achieved by the assessment of relations among predictor and outcome measures through simple correlational and multiple multivariate regression procedures. The correlations were computed to investigate how each separate predictor variable would predict each separate outcome measure. On the basis of the correlational analyses, multiple multivariate analyses were applied to investigate the predictive power of all the predictor variables or subsets thereof for the outcome measures. Because of the lack of studies pursuing similar goals it was found appropriate not to follow a prefixed model for the latter analyses. Alternatively, an explorative working method was applied.

4.3.6.1. PREDICTIONS THROUGH CORRELATIONS

In Table 12 Pearson correlations and corresponding one-tailed probabilities (t-tests) of the predictor with the outcome measures for the total group and both subgroups are reported.

In the total group the correlations of the classical conditioning measures (using auditory, visual and tactile stimuli) with the operant outcome measure are low and not significant, with the two Bayley measures mostly significant and rather high negative. Both operant predictor measures show low positive correlations with the 18 months' operant conditioning measure, and modest negative correlations with the two Bayley Scales. All these correlations are statistically significant. The correlations are generally in the expected direction, that is, positive as far as correlations among learning measures are concerned, and negative for correlations between learning rate measures and Bayley measures. In some instances this trend is not present, but then the correlations are negligible. None of the home-observation measures seem to be strongly related with any of the outcome measures, except the Attention inanimate, which is related quite strongly negatively and significantly with both Bayley measures (see Table 12).

In the high-risk group the relationships are similar to those in the total group. In this group the visual and auditory operant measures are related significantly and relatively strongly with all outcome measures. The correlations are in the expected directions. As shown in Table 12, only the Attention inanimate home-observation variable is related significantly with the two Bayley measures.

The normal group shows only two correlations that are statistically significant and strong, that is, Classical Visual with the 18 months' operant ($\underline{r}$ = .72, $p<.01$) and Operant Auditory with the Bayley Motor ($\underline{r}$ = -.65, $p<.05$). Some other correlations are rather high, but not significant. It has to be noted that the number of subjects in this group is very small.

Table 12
Correlations of the Conditioning and Home Observation Predictor Measures and the Three Outcome Measures

	Operant-18		BayleyMS		BayleyMR	
TOTAL GROUP						
Classical Auditory	.25	(44)	-.36**	(44)	-.46**	(42)
Classical Visual	-.01	(44)	-.16	(44)	-.29*	(42)
Classical Tactile	-.11	(44)	-.52***	(44)	-.54***	(42)
Operant-6 Visual	.31*	(44)	-.47**	(41)	-.38**	(39)
Operant-6 Auditory	.30*	(44)	-.49**	(41)	-.48**	(39)
Attention inanimate	.22	(42)	-.50***	(43)	-.32*	(41)
Attention animate	.15	(42)	-.07	(43)	-.10	(41)
Mouthing	.05	(42)	-.14	(43)	-.19	(41)
Vocalization	-.22	(42)	.07	(43)	-.06	(41)
Manipulation	-.21	(42)	.23	(43)	.10	(41)
Smiling-excitement	.16	(42)	.04	(43)	.00	(41)
Gross movement	-.10	(42)	.20	(43)	.08	(41)
Fussing	.04	(42)	-.02	(43)	.09	(41)
HIGH-RISK GROUP						
Classical Auditory	.18	(33)	-.34*	(35)	-.46**	(33)
Classical Visual	-.15	(33)	-.12	(35)	-.27	(33)
Classical Tactile	-.08	(33)	-.58***	(35)	-.60***	(33)
Operant-6 Visual	.48**	(33)	-.52**	(32)	-.40*	(30)
Operant-6 Auditory	.32*	(33)	-.55***	(32)	-.47**	(30)
Attention inanimate	.26	(31)	-.52*	(34)	-.39*	(32)
Attention animate	.16	(31)	-.03	(34)	-.12	(32)
Mouthing	-.02	(31)	-.09	(34)	-.19	(32)
Vocalization	-.27	(31)	.12	(34)	-.01	(32)
Manipulation	-.23	(31)	.21	(34)	.16	(32)
Smiling-excitement	.19	(31)	.05	(34)	.03	(32)
Gross movement	-.16	(31)	.18	(34)	.05	(32)
Fussing	.03	(31)	.04	(34)	.12	(32)
NORMAL GROUP						
Classical Auditory	.39	(11)	-.05	(9)	-.23	(9)
Classical Visual	.72**	(11)	-.05	(9)	-.07	(9)
Classical Tactile	.18	(11)	-.14	(9)	-.25	(9)

Continuation table 12

	Operant-18		BayleyMS		BayleyMR	
Operant-6 Visual	-.39	(11)	.13	(9)	-.07	(9)
Operant-6 Auditory	.18	(11)	.38	(9)	-.65*	(9)
Attention inanimate	.09	(11)	-.45	(9)	.38	(9)
Attention animate	.11	(11)	.03	(9)	.38	(9)
Mouthing	.16	(11)	-.34	(9)	.13	(9)
Vocalization	-.07	(11)	-.21	(9)	-.46	(9)
Manipulation	-.17	(11)	.17	(9)	-.54	(9)
Smiling excitement	.08	(11)	-.11	(9)	-.32	(9)
Gross movement	.04	(11)	.38	(9)	.30	(9)
Fussing	.05	(11)	-.39	(9)	.32	(9)

Note Figures in parentheses indicate the number of subjects entered into the analysis.
*p < 05, **p < 01, ***p < 001, all one-tailed

In summary, several of the early measures seem to explain some variance in the outcome measures. The strongest predictors of both Bayley Scales and the Operant-18 measure, at least for the total and high-risk group, seem to be the Classical Auditory and Tactile measures, the Operant-6 measures and in addition the Attention inanimate variable (home-observation). The signs of the correlations among these early learning rate measures and all predictor measures are as hypothesized: they are positive when relations between learning rate measures are concerned, and negative when learning rate measures are related with Bayley scores.[5] The observed negative correlations among the Attention inanimate and the two Bayley outcome measures is striking. Most of the significant correlations found are higher in the high-risk than in the total group. The classical measures show quite consistently higher correlations with the Bayley Motor than with the Bayley Mental. For the Operant-6 measures this latter tendency is in the opposite direction. In the normal subgroup the Classical Auditory, and notably the Classical Visual, but also the Operant-6 are correlated quite strongly with the Operant-18 variable. In addition, the Operant-6 Auditory variable is related strongly with the Bayley Motor scores. These latter correlations cannot easily be interpreted, although a stronger relationship seems to exist between the classical measures and the 18 months' operant in this subgroup than in the high-risk subgroup. From these results it cannot be concluded whether the various correlations account for unique or overlapping variances in the outcome variables. Such information can be obtained by performing multiple regression analyses, the results of which will be reported in the following section.

5 When rank-ordering the scores (allotting the highest rank to the highest scoring or the 'unsuccessful' subjects), and computing Spearman rho-correlations, similar results appear.

4.3.6.2. MULTIPLE UNI- AND MULTIVARIATE ANALYSES

In the previous section, some of the applied predictor measures had significant correlations with the applied outcome measures. In this section, analyses among predictor and outcome measures were performed by means of univariate and multivariate multiple regression analyses, applying F-tests and Hotelling's T-squared statistics.[6] Their results are shown hereafter for the total experimental group and for the high-risk subgroup. For the normal subgroup analyses were found not useful (and thus were not performed) because of the small sample size ($\underline{n} \leq 12$) and associated loss of degrees of freedom in multiple analyses. Subjects with allotted scores are included in the analyses.

To get more insight in the strengths and nature of the relationships among the sets of variables, the analyses were carried out in a sequence as outlined in Table 13. This sequence was determined by the main goal of this explorative study: exploring the predictive strength of classical and operant conditioning measures for later learning (assessed by an operant conditioning procedure) and developmental level (assessed by the Bayley Scales of Infant Development). An extension of the aforementioned goal was to develop an assessment battery (aiming at prediction of future development) which would be as powerful and as economic as possible.

First, the results of the analysis with the set of three classical conditioning measures as independent variables and the set of 6-months' operant conditioning measures (as dependent variables) are reported in *ANALYSIS I*. This is the only analysis where not the 18-months' operant conditioning and both Bayley measures were used as dependent variables, but the two Operant-6 measures. This analysis was performed to assess whether at an age as early as 4 months and by the use of classical conditioning measures it is possible to predict learning rate measures taken at the age of 6 months (Operant-6 measures). Secondly, the relationship between the same set of classical conditioning measures (independent variables) and the 18-months' outcome measures (dependent variables) was established. Then, the same analyses were performed with the set of Operant-6 measures alone, and the set of observation measures alone as independent variables (*ANALYSES II to IV*). These analyses were executed to assess the predictive power of each of the aforementioned sets of variables separately. After the aforementioned investigations, including single sets of independent variables, combinations of these independent sets of variables were formed to carry out the analysis (*ANALYSES V*

6 Because normality over groups is a prerequisite condition for all aforementioned regression analyses this was tested. Indeed, several of the measures used showed skewed distributions, deviating statistically significant from normal distributions (Kolmogorov-Smirnov tests, not reported here). Both for the total group as well as for the subgroup of high-risk subjects normality of the following variables was rejected: all classical conditioning variables, BayleyMR, Gross movements, Crying-fussing (the latter two are home observation measures) In addition, the Operant-6 Visual and Vocalization (a home observation variable) showed a skewed distribution for the total sample of subjects Although violation against the aforementioned conditions of normality of score-distributions seems relatively unimportant for regression analysis because of reasons mentioned by Winer (1971) and Norusis (1986), logarithmic transformations on the scores of the variables concerned turned out to be useful (as suggested by Winer (1971) and Kirk (1982)). These latter transformations indeed led to distributions which did not differ significantly from normal distributions (Kolmogorov-Smirnov tests) The results of multivariate multiple regression analyses on the data after such transformations (and applying identical variables as in the analyses reported in this chapter), showed not to be essentially different from results of analyses with untransformed data Therefore, we decided to show the regression analyses on the original data

and VI), as presented in Table 13. In this way, the additional prediction provided by each set of predictors to the other sets could be assessed (Jay & Farran, 1981). In addition, a combination of three selected variables was made to enter into the analysis (*ANALYSIS VII*). As already mentioned above, a schematic presentation of the successive data analysis sequence, as described in this section, is shown in Table 13.

Table 13
Sequence in the Regression Analyses* and, in Parentheses, the Total Number of Predictor Variables Entered. For all Analyses the Three 18-Months' Outcome Measures are Included as Dependent Measures, Except for the First Analysis in which the Two 6-Months' Operant Measures Were Used

SEQUENCE	SETS OF PREDICTORS	DEPENDENT MEASURES
I	All Classical measures (3)	Operant-6 measures (2)
II	All Classical (Class) measures (3)	Operant-18 BayleyMS/MR
III	All Operant-6 measures (2)	Operant-18 BayleyMS/MR
IV	Observation (Obs) measures (8)	Operant-18 BayleyMS/MR
V	Class + Operant-6 measures (5)	Operant-18 BayleyMS/MR
VI	Class + Operant-6 + Obs measures (13)	Operant-18 BayleyMS/MR
VII	Class Auditory + Operant-6 Auditory + Attention inanimate (3)	Operant-18 BayleyMS/MR

Note. Class = Classical; Obs = Observation; Att = Attention
* The regression analyses were executed for the total group and the high-risk group, but not separately for the normal group (small sample size).

The results of the aforementioned analyses are summarized in Tables 14 and 15. In these tables the accounted variances are presented under the heading 'Adjusted R-squared', which points to adjustment of the squared multiple correlation (accounted variance) associated with the loss of degrees of freedom (d.f.) due to multiple comparisons.

ANALYSIS I: All classical measures as predictors of the Operant-6 measures. The analysis with the first set of variables, that is, the three classical conditioning measures, as predictors of the two 6-months' operant measures (Operant-6 Visual and Operant-6 Auditory) revealed that the set of outcome measures was not significantly predicted (Hotelling's criterion), neither for the total group of subjects, nor for the group of high-risks (see Table 14). A significant multiple correlation was found with the Operant-6 Visual measure in the high-risk group. These results indicate more resemblance in scores between the classical tasks and the Operant-6 Visual task in the high-risk group.

Table 14
Regression Analyses of Prediction of the Operant-6 Variables by the Classical Conditioning Measures, for the Total Group and High-risk Group

Set with outcome variables	Adjusted R-squared	F-value (d.f.)	Sign. level
Total group (N = 45)			
Operant-6 Visual	.07	2.16(3,41)	.11
Operant-6 Auditory	.04	1.64(3,41)	.20
High-risk group (n = 33)			
Operant-6 Visual	.16	3.04(3,29)	.05
Operant-6 Auditory	.01	1.07(3,29)	.38

Note. Hotelling's T-squared not significant.

ANALYSIS II: All classical measures as predictors of the 18 months' measures. The same analysis, but applying variables as outlined in *ANALYSIS II* (Table 13), shows that the multivariate predictions are significant, employing Hotelling's T-squared (see Table 15). Variation in the classical conditioning variables does explain significant portions of the variance of the 18-months' variables combined. Moreover, multiple regression analyses show significance in both groups. Accounted (adjusted) variances may be regarded as modest. The variables that predict a significant portion of the outcome measures are the Classical Auditory and Classical Tactile measures (t-tests, not reported in a table).

ANALYSIS III: Operant-6 predictor measures. The multiple regression analyses with the 6-months' operant conditioning measures as predictors, and the 18-months' operant as well as both Bayley measures as dependent variables, indicate significant relationships in both groups (see Table 15). When looking at the outcome measures separately, it can be seen that the predictors are related with the 18-months' operant measure, as well as – more strongly – with the two Bayley measures. Percentages of between 14 and 37 of the variances (adjusted for the number of independent variables) in the outcome measures is explained by the two operant predictors, which could be considered modest. In addition, it could be concluded that the relationships are somewhat more pronounced in the high-risk group because the significance of the applied multivariate test (Hotelling's T-squared) is of the same level (despite the smaller sample size). The accounted variance levels are comparable with those found in ANALYSIS II, in which the classical measures were applied. In both groups the Operant-6 Auditory conditioning measure shows to be the more powerful predictor of the Bayley outcome measures (not reported in a table). The strongest effect is shown on the Bayley Mental measure. The operant conditioning outcome measure is related strongest with the Operant-6 Visual variable (not reported in a table).

ANALYSIS IV: Observation predictor measures. As can be seen in Table 15 the set of observation variables does not show a significant effect for the combined outcome measures. This holds true for both the total group as well as for the subgroup of high-risk infants. When considering the outcome measures separately, one could conclude that both Bayley measures are related stronger to the observation variables than the Operant-18 measure (see the results of the F-tests). The magnitudes of the predicted variances are comparable with the two previously investigated sets of predictor variables. The number of variables entered, however, is quite a bit higher, possibly leading to the observed accounted variances. In addition, when considering the predictor variables, it shows that particularly the Attention inanimate, and to a lesser extent the Manipulation and Attention animate measures contribute considerably to the variance explained in the two Bayley measures. In general the Bayley Mental Scale measure seems to have the strongest relationship with these latter predictor variables (these latter results are not reported in a Table).

Combining the sets of predictors will provide information on the distribution of unique portions of variance over the different predictors. The next analyses are directed at the latter question.

ANALYSIS V: Classical + Operant-6 predictor measures. As shown in Table 15 large and significant portions of the variance of the outcome measures is explained by the two sets of conditioning measures (Classical and Operant-6) for both groups (the total group and high-risk subgroup). That is between 19 and 44% (adjusted) variance is explained in the total group, and between 44 and 56% in the high-risk subgroup. All three outcome measures are significantly influenced, but the BayleyMS undergoes the greatest influence (as can be seen in the results of the F-tests). Further analysis shows that when applying single dependent variables the Operant-18 is explained most by the Classical Auditory and Classical Visual variables, while both Bayley measures are related fairly strongly with the Operant-6 Auditory and Classical Tactile measures (not reported in a Table). In short, the relationships found in the total group are also found in the subgroup of high-risk infants, but the latter relationships show to be more pronounced.

ANALYSIS VI: Classical + Operant-6 + Observation predictor measures. The number of predictor variables entered into the analysis is again increased. Although the maximum number of available predictors is used to predict the three outcome measures, and the variances explained are mostly higher (but degrees of freedom are lost), most relationships are still significant (see Table 15). Particularly, the accounted variance for the BayleyMS is increased (in comparison with the previous analysis), due to the set of observation variables.

Preliminary conclusions on the basis of all the multiple regression analyses presented thusfar could be the following. First, most multiple predictions are statistically significant. These relationships seem to exist in the total group, but being more specific by excluding the 'to be normal' infants, those predictions even gain in prominence in the high-risk group.

Table 15
Regression Analyses of Prediction of the Outcome Variables by Different Sets of Predictor Variables, Indicated by Roman Numerals Which Correspond to Those Used in Table 13. Asterisks Indicate the Different Levels of Significance of F-tests for the Separate Outcome Measures

Sets of Predictors		Adjusted R-squared Operant-18	BayleyMS	BayleyMR	Significance Hotelling's T-squared
II	All classical measures				
	Total group (39)	.09	.29**	.33**	.004
	High-risk (30)	.25*	.36**	.39**	.001
III	Operant-6 measures				
	Total group (39)	.14*	.30**	.24**	.008
	High-risk (30)	.22*	.37**	.22*	.008
IV	Observation measures				
	Total group (38)	.02	.34**	.21	.18
	High-risk (29)	.09	.32*	.32*	.25
V	Classical + Operant-6 measures				
	Total group (39)	.19*	.44***	.42***	.001
	High-risk (30)	.44**	.56***	.48**	.000
VI	Classical + Operant-6 + Observation measures				
	Total group (38)	.14	.63***	.42**	.004
	High-risk (29)	.55**	.65**	.46*	.007
VII	Classical Auditory + Operant-6 Auditory + Attention inanimate				
	Total group (38)	.14*	.60***	.43***	.000
	High-risk (29)	.19*	.59***	.42**	.000

Note. Figures in parentheses indicate the number of subjects entered into the pertaining analyses.
*p<.05, **p<.01, ***p<.001.

Secondly, a clearer idea has been developed about the kind of measures which play significant roles in the relationships and, by the same token, which do not. From the analyses it could be concluded that unique portions of the variance of the 18-month operant conditioning measure are explained by the classical and operant learning measures, while substantial portions of the Bayley Scales are explained by the learning and the observation measures. More insight in the specific predictor variables that account for the variances can be gained by looking at the standardized regression coefficients (not reported here). Inspection of the latter analyses reveals the following. The conditioning measures applying auditory stimuli seem to have special relevance as predictors. Moreover, the Attention inanimate and Attention animate observation variables also play a role of importance in the regression analyses. To a lesser extent, the following variables seem powerful: Classical Tactile, Operant Visual, and Manipulation.

Although the aforementioned accounted variances are reasonably high (and relationships are statistically significant), the number of variables on which they are based is great. An attempt to reduce the number of predictor variables is to exclude those predictor variables which, on the basis of the analyses applied, have shown to have weak relationships with the outcome variables. The next analysis investigates these latter considerations.

ANALYSIS VII: Classical Auditory + Operant-6 Auditory + Attention inanimate predictor measures. This analysis was performed to find out what effect a reduction in the number of predictor variables entered would have on the variance explained in the outcome measures. This was done to investigate on and increase the practicality of the applied procedures by determining which variables (as few as possible) should be applied while maintaining optimal or sufficient clinical relevance. Only one of each set of conditioning measures and one attention observation measure in a multivariate analysis (a total of three variables) was applied. The latter measures were selected on the basis of the already presented analyses, which had shown them to have the strongest multiple correlations with the three outcome measures.

As shown in Table 15, for the total experimental group Hotelling's T is highly significant ($p < .000$), while the variances explained are 14%, 60%, and 43% for the three outcome variables. The aforementioned results are similar for the high-risk subgroup. Thus, excluding another five predictors (out of a total of eight) resulted in a relatively small reduction of explained variance for the three outcome measures, compared with the previous step in the two groups. From a practical point of view, this drastic reduction in the number of measures and a relatively minor reduction in predictive power seems relevant.

4.3.6.3. SUMMARY OF CORRELATIONAL AND MULTIVARIATE PREDICTIONS

In general, the relationships found for the high-risk subgroup are similar to those for the total group (a substantial portion of the subjects in the total group form the high-risk group). However, when using all the early conditioning measures (classical and operant) to predict outcome measures, a substantively greater part of the variance of the outcome measures is explained in the high-risk group (compared with the total group), and the statistical significance is increased. Of all the measures applied, the Classical Auditory and Operant-6 conditioning predictors, together with the Attention inanimate observation measure, seemed to be the most powerful predictors of the outcome measures (regarding the variances accounted for). For the total group, and even more for the high-risk subgroup, about half of the outcome measures' variances were accounted for by the predictor measures (the conditioning and observation measures). Applying only the 4 months' Classical Auditory, the 6-months' Operant Auditory and the Attention inanimate measures resulted in accounted variances of between 14% and 60% for the total group and between 19% and 59% for the high-risk group (variances adjusted for the number of dependent variables). Of the outcome measures, the explained variances differ considerably, that is, the operant measure shows the lowest, and the Bayley Mental measure the highest accounted variance.

In summary, the conditioning variables, notably the combination of the Classical Auditory (at 4 months), the Operant-6 Auditory and Attention inanimate observation measure accounted for a substantial, and thus meaningful, portion of the variances in the outcome variables, notably the scores on the Bayley Scales. The combined predictor variables showed stronger relationships with the Bayley Mental scores than with the Motor Scale scores. In the high-risk group the combined conditioning scores accumulated a high percentage of accounted variance in the operant conditioning outcome measure.

4.3.7. Comparison of successful and unsuccessful high-risk subjects

The results of the analyses as presented in the previous section point to a rather strong predictive relationship of the Classical Auditory, the 6-months' Operant Auditory conditioning procedures and the Attention inanimate observation measure. These findings were based on the use of raw (for all measures) and allotted scores for some of the conditioning procedures in part of the subjects (conditioning failures). Furthermore, a rationale was given for the use of these allotted scores.

In this section, the meaningfulness of successfulness versus unsuccessfulness in the conditioning procedures as a marker variable is investigated through screening into the two categories (Allen & Yen, 1979) and relating them to later developmental level (Bayley scores). This procedure could provide data-based support for the inclusion of the unsuccessful subjects (and the use of allotted scores) and for the clinical significance thereof. Following the aforementioned reasoning, groups of 'successful' and 'unsuccessful' subjects of the subgroup of high-risk infants were composed. Then the groups' scores on the Bayley measures were computed. For the comparison only the group of high-risks is used, because of a presumed group-homogeneity and because proportionally more unsuccessful subjects were found in this group than in the subgroup of normals.

Table 16
Average Scores of 'Successful' and 'Unsuccessful' High-risk Subjects

	Means		t-Test
	Unsuccessful	Successful	Significance
BayleyMS	107.3 (12)	123.8 (20)	$p < .01$
BayleyMR	44.2 (11)	51.5 (19)	$p < .05$

Note. Figures in parentheses indicate the number of subjects participating in the analyses.

In Table 16 the mean scores on the two Bayley measures are shown for both groups. Inspection of Table 16 shows that the differences between the means for the groups on the variables are statistically significant.[7] The direction of these differences in general points to a disadvantage for the 'unsuccessful' infants.

Tentatively, one might conclude from these analyses that the conclusion of Bathurst and Gottfried (1987), that "missing data resulting from unresponsiveness or lack of cooperation to standardized testing conditions may serve as a marker variable" is confirmed. In addition, the inability to complete testing procedures (for as yet unknown reasons) could be considered such a variable. Thus, untestability or inability to learn (within a limited but 'ample' number of training trials) could have developmental implications. Inclusion of the 'unsuccessful' subjects, and consequently, allotment of scores in this study, seems to have been a viable procedure.

4.3.8. Classification of all subjects by discriminant analysis

Discriminant Analysis enables one to investigate the viability of prediction through classification of individual subjects. For this classification all the subjects (the high-risk and normal infants) were divided over three, and finally two, groups according to the individual performance on the Bayley Mental Scale (grouping variable), applied at the age of 18 months (cf. Drotar & Sturm, 1988). The choice of the grouping variable was based on the argument that the Bayley Scales comprised the main outcome measure and can be considered a reliable instrument for developmental assessment. However, some of the infants turned out to have motoric problems (affecting the Bayley Motor Scale score). This made the Bayley Motor Scale (or the composite Bayley Mental and Motor Scales scores) less suitable as a classification variable, although in this study the correlation between the Mental and Motor Scale scores was high. In addition, it was assumed that the Bayley Mental Scale would comprise a more valid index of mental development (see also Rubin & Balow, 1979).

Inclusion in the low-functioning group was based on a Bayley Mental raw score of 112 or less (cutting score). The high-functioning group included the subjects who scored 139 or more. The middle group was made up of the remaining infants (scores between 113 and 138). In identical analyses the subjects of the middle and high group were combined, comprising a group of average or above average scorers on the Bayley Mental Scale. Four infants (one high-risk and three normals) could not be grouped because of missing scores on the grouping variable. On six subjects one or more discriminating variables were missing. All missing scores resulted in a total of eight subjects who could not undergo the analyses. The choice of the aforementioned cutting scores was based on the rounded mean (126) and standard deviation (6.41) of the raw scores on the BOS 2-30 standardization sample of normal 18-months-old children, that is, subjects were included in the high or low functioning group when they differed more than two standard deviations from this latter mean.

The discriminating variables, that is, Classical Auditory, Operant-6 Auditory, and Attention inanimate, were a selection of the predictor variables. This selection was based

---

7 The same analyses were performed applying rank-ordered scores (Mann-Whitney U test). This was done to let the statistical analyses be in accordance with the rather gross criterion applied for group-membership, that is, successfulness. The results of these analyses point to a significant difference in means between the two groups on the BayleyMS measure.

on their predictive power as shown in the correlation and multivariate multiple regression analyses (see the appropriate results sections). The analyses were performed by applying two and three groups to see if classification in these groups would be meaningful with regard to discrimination of higher, average, and lower functioning subjects or discrimination of lower and average and above subjects.

The resulting discriminant functions were significant. The analyses show that 82.5% of the grouped cases was correctly predicted in the three groups (hit rate), while a slightly higher percentage, that is, 87.5% of the cases, was correctly predicted in the two groups procedure (low group and the combination group) (see Tables 17 and 18). When applying the three- as well as the two-groups procedure, the results show a 100% prediction accuracy for the high-functioning subjects (while the number of subjects is very small). Prediction accuracy was high, that is, a percentage of 80 or more, for the other subjects.

Table 17
Classification of Subjects, Applying the Classical Auditory, Operant-6 Auditory, and Attention Inanimate Measures as Discriminating Variables, in the Three Groups Procedure

Actual Group	No. of Cases	Predicted Group Membership Low	Middle	High
Low	7	6 (85.7)	1 (14.3)	0 (0)
Middle	31	3 (9.7)	25 (80.6)	3 (9.7)
High	2	0 (0)	0 (0)	2 (100)

Note. Total percentage of "grouped" cases correctly classified: 82.5%.
Figures in parentheses are the percentage of grouped cases correctly classified.

Table 18
Classification of Subjects, Applying the Classical Auditory, Operant-6 Auditory, and Attention Inanimate Measures as Discriminating Variables, in the Two Groups Procedure

Actual Group	No. of Cases	Predicted Group Membership Low	Combined
Low	7	6 (85.7)	1 (14.3)
Combined	33	4 (12.1)	29 (87.9)

Note. Total percentage of "grouped" cases correctly classified: 87.5%.
Figures in parentheses are the percentage of grouped cases correctly classified.

From the analyses it is clear that the hit rates per predicted group were high. In addition, one could conclude that from the subjects who were, according to the scores on the Bayley Mental Scale, on the lower side of the developmental continuum, one (out of 7) subject was predicted (according to the scores on the applied predictor variables) in the middle group (false negative) and none in the high-scoring group. Of all the subjects that were predicted in the low group three or four subjects out of nine or ten were in the middle group (or higher) (false positives, defined as those subjects that are predicted as developmental risks but actually show a favorable outcome). A rather high proportion of false positives could be preferred to a high number of false negatives, because the negative consequences of a false positive may not be as great as the concequences of a false negative decision (Siegel, 1982).

4.4. Summary and discussion

Classical and operant conditioning strategies for the assessment of learning rate and home observation measures were used in this longitudinal study for prediction of later psychological development, measured by an operant conditioning procedure and the Bayley Scales of Infant Development. Two groups of subjects were involved, that is, a small group of normal infants and a group of high-risk and handicapped infants. The investigations took place when the subjects were 4, 6, and 18 months old. At the latter age the outcome measures were taken.

On all predictor as well as outcome measures, not including the home observation measures, the mean scores of the high-risk subgroup were at a disadvantage in comparison with the normal group. The latter differences were never, except for the Bayley Mental Scale scores, significant. The results of simple correlational and multiple multivariate regression analyses indicate that the scores on two of a total of three of the classical and both the operant procedures (predictor measures) accounted for a substantial amount of variance in the outcome measures, that is, particularly the scores on the Bayley Scales. Both operant predictor measures showed moderate significant correlations with the operant outcome measure. The home observation variables used in this study were not related to outcome variables, except the variable 'Attention inanimate'. The classical conditioning procedures showed to be easily applicable, that is, only one conditioning failure was observed. However, the operant conditioning procedures were much more vulnerable to failure. More particularly, almost 6% of all the conditioning occasions was unsuccessful. The conditioning failures were not caused by obvious state problems (sleeping, fussing or crying). Furthermore, from a comparison of high-risk children who conditioned successfully within the maximum number of training trials allowed, with those who did not (conditioning failures), it was concluded, that to maintain instead of to remove these 'unsuccessful' subjects from the study samples was meaningful in view of the goal of this study. In addition, classification of subjects on the basis of the three most powerful predictor variables (by means of a Discriminant Analysis procedure) resulted in a reasonably high percentage of correct group-predictions, that is, more than 80 percent. This latter result points to the possible use of the aplied measures for the benefit of individual predictions. In the following subsections several considerations are made regarding the data, statistical procedures and results of the present study. Moreover, these findings will be related to other research evidence and their possible implications will be discused.

4.4.1. Data preparation and analysis

Because of the heterogeneity of the sample of high-risk and handicapped subjects (and often associated greater dispersion of scores, cf. Rose et al., 1989) and the nature of the applied measures (learning rate by conditioning procedures) normal distributions of the scores and homogeneity of the variances, as estimated from the samples, could hardly be expected. Although the applied data analysis procedures are considered to be robust against violations of the aforementioned conditions, in many instances non-parametric data-analysis procedures were performed parallel to parametric procedures. These non-parametric procedures, however, yielded similar or identical results. Thus, the procedure followed contained sufficient safety to present the results of parametric statistical procedures on the raw scores.

The use of allotted scores in the present study seems to be a valid approach in view of the following. First, subject loss due to inability to condition within a preset number of training trials, while the infants had not showed state problems, would in all probability have resulted in biased findings. That is, in several studies 'lost subjects' showed to distinguish in important criterion measures (cf. Bathurst & Gottfried, 1987). Secondly, as an informal control, the data were also analyzed by non-parametric tests, applying rank-ordered scores, thereby providing the unsuccessful subjects (conditioning failures) the highest ranks. This procedure overcomes to a large extent possible drawbacks of the use of allotted scores. Moreover, the latter analyses led to similar results as with the parametric tests. Thirdly, the scores of the successful and unsuccessful subjects (the latter group implied allottment of scores) were analyzed separately, showing significant differences in mean scores on the Bayley Mental Scale between the two groups, the successful subjects being in advantage, a finding which is in agreement with those of Bathurst and Gottfried (1987). The latter results point to the need of further in debt investigation of subject loss in studies where active participation of the subjects is required (Wachs & Smitherman, 1985). Moreover, procedures to overcome problems with data analysis, like the one suggested and used in this study, are needed badly. Data adaptation procedures, similar to the aforementioned, are not common. However, in a predictive study, using high-risk very preterm infants, semi-arbitrary scores for standardized as well as nonstandardized tests were assigned in case of very disabled children and in case of refusals to perform specific test items (Costello et al., 1988). Not only on behalf of research could the aforementioned problem and its solution be considered relevant, but also in clinical practice, where assessment of difficult-to-test children is an area of much interest (cf. Van Hasselt & Hersen, 1987).

4.4.2. Predictors: Learning assessments and observation

The data analyses reveal that with regard to the research questions the classical conditioning procedures could be considered successful, with simple correlations comparable to or higher than those reported for example for the HOME and attention measures. For the latter measures correlations of .30 to slightly more than .50 were reported in samples of normal as well as high-risk infants and young children (e.g., Fagan & Singer, 1983, Lewis & Brooks-Gunn, 1981, and Ruddy & Bornstein, 1982, among others). When applying the HOME measures, correlations of .20 to .50 have been reported in medically normal samples of which some participated in long-term investigations because of social-economical disadvantage (see section 3.3, and Bradley & Caldwell, 1980, 1984a, 1984b; Bradley et al., 1989). The findings in the latter studies as well as in the present study point to similar predictive values. The question is, whether this agreement in findings is related to similarities in subject-characteristics tapped by the procedures. This could be assumed for the attention and conditioning, and perhaps for the HOME measures. That is, these measures require sustained and/or repeated attention for the presented stimuli and memory and discrimination abilities for differential responding.

Regarding the present study, particularly the procedure with the auditory CS when it is used in multiple regression analyses in combination with other selected variables seems significant as a predictor. This measure shows a statistically significant (but low) correlation with both Bayley measures and shows in combination with other predictor measures (i.e., the operant predictors) to account for significant portions of variance in the outcome measures (more than 40% in both Bayley measures). Therefore it could be concluded that the applied procedures, or classical procedures in general, applied at this age, tap functions or systems (abilities) that are related directly with the measures of general mental ability taken at a later age.

These findings are also in accord with the contention and some empirical evidence that classical conditioning is related with the maturity or condition of the neurological system. For example, Ross and collaborators (Ross et al., 1967) found, when applying two groups of Down's syndrome children of different ages, a relation with age. That is, the older group reached much higher levels of conditioning. In a recently reported study, Hoffman, Cohen, and DeVido (1985) compared classical eyelid conditioning in adults and 5- to 14-months-old infants. They found similar results as in the present study, while applying a somewhat different conditioning procedure. That is, no conditioning failures occurred and the infants learned slower and with more variability in scores than the adults. Moreover, the data reported in the Hoffman et al. study were transformed into learning rates by the present author. On the basis of the latter transformations, it could be prudently maintained that the learning rates found seemed comparable to the ones found in the present study.

Several intra-individual factors could possibly influence predictive relationships. For example, the adverse effects of a possible neurological involvement in the high-risk and handicapped population may become manifest at a later age (Vlugt, 1979) when cognitive functions become progressively more prominent. It may be argued that the significance of this latter observation depends heavily on the sensitivity and validity of the applied assessment procedure. Suggestions for recovery of organically based defects, with effects in the opposite – advantageous – direction, are made as well. An indication for the viability of the aforementioned delay in actualization of adverse effects could be considered to be the observation that several of the subjects who showed deficits in development at 18 months of age had conditioned rather fast at the age of 4 months. Of course it is possible that the classical conditioning procedures were also successful with subjects with diminished CNS integrity, or in other words, the neural mechanisms which are required for successful classical conditioning may need to be quite limited, a suggestion made by several authors (Papousek, 1969; Tuber, Berntson, Bachman, & Allen, 1980). The present study provides some evidence for the assumption that classical conditioning procedures are, at least to a certain extent, sensitive to neurological involvement and/or its long-term effects. The same reasoning also holds for the applied operant conditioning procedures.

With respect to the sensitivity of the applied classical conditioning procedures, regarding detection of children at higher risk of delays in development due to non-optimal learning processes, a possible source of improvements is offered by the way the responses are measured. In fact, it has been suggested that characteristics of the response and acquisition of it need finer measurement methods, for example, (neuro)-physiological, for detection of more fundamental features of the learning process. Classical conditioning procedures seem to satisfy in particular the requirements for such measurement methods (Dawson & Schell, 1987). Application of such measurements, however, would place restrictions on their practicality.

The classical conditioning procedures have shown to be applicable successfully with children of this age and biological status. That is, not only in this research – as could be concluded by the high percentage of infants that succeeded conditioning – but also in research aimed at other research-questions (Hoffman et al. 1985; Lancioni et al, 1985, 1989). This shows these procedures to be rather resistent to subject loss due to state conditions and the like in very young infants. Subject loss is experienced quite often in basic as well as applied infancy research and considered a serious problem (Wachs & Smitherman, 1985). Furthermore, these procedures are applied in a highly standardized way. Application of existent standardized assessment procedures, aimed at the evaluation of infant developmental status, among other things, seems to be rather problematic in very young normal infants. In high-risk and handicapped populations, in which such assessments are in bad need, this problem is even more pronounced. Classical conditioning procedures seem on the one hand to satisfy several of the aforementioned needs, and on the other hand to be able to overcome some of the problems.

In addition to the classical conditioning procedures the operant conditioning predictor measures hold considerable promise for the prediction of later psychological outcome. The observed correlations with later developmental measures could be explained by the assumption that operant conditioning may tap, or at least has significance for, cognitive abilities, similar to those the Bayley Scales appeal to or precursors thereof. That is, operant conditioning procedures may tap higher mechanisms of the CNS (and cognitive system) and, in addition, may make use of available environmental contingency experiences, while also other factors, such as motivation and temperament, play a pertinent role. The operant procedures are then more vulnerable with respect to experimental control (shown by the number of conditioning failures, which may reflect actual deficits in learning). At the same time they may tap abilities and incorporate experiences that are more similar to those used and needed for learning and adaptation to everyday life circumstances. Thus, operant conditioning procedures may imply the incorporation of several learning variables, such as, neural integrity, attention, arousal, selection of salient stimuli or relevant features from the environment, contingency experiences and storage of them. An important issue for investigation would then seem to be the search for the latter suggested learning variables and the way they exert their influence on the performance on concurrent classical as well as operant tasks. Moreover, their long-term significance could be pursued (cf. Gaussen, 1984).

With respect to the applied operant conditioning procedures, the study of Fagen et al., (1987) seems to be of special relevance as a reference point in this specific area. They used free operant conditioning procedures and retention – 7 or 14 days – of learned responses as the most important measures, applied at the ages of 3, 7, and 11 months, to predict performance on the Stanford-Binet Intelligence Test and the Bayley Scales of Infant Development at 27 months of age. The data of their research point to significant correlations with an average of .40 between the retention measures and the Stanford-Binet and Bayley Mental Scale scores. The authors maintain that infant memory is a relatively stable component of infant cognition and that measures of infant memory should be incorporated in screening devices for infants at risk for retarded mental development. The data of the Fagen et al.'s study are consistent with the correlations found between the 6-months' operant learning rate measures and the Bayley Mental Scale scores, as well as the 18-months' operant conditioning scores, in the present study.

With regard to the home-environmental measures, one of the variables, that is, the Attention inanimate variable, correlated significantly with the outcome variables and added to accounted variance when used in combination with the conditioning predictors. The observation data, however, were taken in a rather limited period. It could be argued, that if more data had been available at this level (i.e., collected in more situations, thus representing more accurately the infants' interactions with the environment), more definite conclusions could be drawn.

Individual prediction (by means of discriminant analysis) from the assessments in the first half year of life to the age of 1 1/2 years, resulted in a substantial number of children being correctly classified (more than 80%, or an error rate of less than 20%). Thus, error rate could be considered rather low. However, the classifications were based on rather gross criteria. Error rates in other studies that have attempted identification in the first year has varied from 6% to close to 50% depending on the subject population, the measures, and the age to which prediction was made (Cohen & Parmelee, 1983; Field, Hallock, Ting, Dempsey, Dabiri, & Shuman, 1978; Hunt, 1981; Siegel, 1982).

4.4.3. Outcome measures: Learning assessment and Bayley Scales

The 18 months' operant conditioning strategy showed moderate but significant correlations with both early operant measures, particularly in the high-risk sample, suggesting long-term stability of learning rates measured by similar but age-adapted procedures. In addition, the outcome conditioning variable showed moderate significant correlations with both Bayley Scales. Particularly the Mental Scale presented consistent correlations of about .54. These results demonstrate relationships between both previous and concurrent measures of learning and assessment of general cognitive functioning.

In this study we have come across a rather high percentage of conditioning failures, not only in the high-risk and handicapped sample, but also in the supposedly normal children. This latter finding makes the measure not suited as a sole measure of developmental level and in addition one could question its validity. Data on the long-term significance of the criterion of successfulness of this 18-months' conditioning measure are not available. Similar reasons for the high proportion of conditioning failures could be suggested as those pertaining for the already mentioned 6 months' operant procedures. Informal observations suggest that several factors accounted for the procedural problems for different children. That is, for some children the apparatus (and provided contingencies) did not seem sufficiently appealing (from the start or during a session); other children seemed not to be able to regulate their attention to the required distinctive response, in other words, distractions that may be internal or external to the infant. Although speculative, it could be maintained that these observed behaviors (and the variables which control these behaviors) also have an impact on everyday learning. In that sense these variables should also play a role in devising individual intervention programs (Wachs, 1984; Wallace, 1988). Assessment of infants, particularly when they are handicapped, seems to require procedures and stimulus-characteristics that are adapted to specific groups of children, or even individuals. Moreover, assessment of learning should preferably incorporate more test occasions in order to be less dependent from possible temporary state instability or the like.

4.4.4. Conclusion

In conclusion, it is hard to compare the results of this study with those of other studies, pursuing the same goal. Moreover, most studies cannot easily be compared because of the different subjects and ages used, in addition to the types and numbers of measures. Nevertheless, our findings and those of many others who applied neurophysiological, environmental (HOME), and attention measures at a very young age, point to a slightly but significant better prediction of future development than achieved with infant tests, particularly with regard to long-term prediction for some of the aforementioned measures. The development of a screening battery composed of developmental (e.g., learning, attention, memory), neuro(physio)logical and environmental measures (e.g., by observations in a natural situation) for detection of later developmental disability would seem relevant.

When comparing the results of the present study with those of attention measurements, one could conclude that these are in a similar range when the measures of each set (i.e., the sets of classical and operant conditioning and observation predictor measures) are considered separately. However, the results of this study show the power of combinations of predictors. Moreover, when a selection of measures from the different sets was made (i.e., the most powerful predictors from learning as well as from observational (environmental) measures) rather high proportions in the outcome measures' variance were accounted for. It should, however, be realized that the data were taken from a rather heterogeneous sample, as compared to most other studies, which could also have contributed to the higher predictive values found. In addition, most study samples, as well as the present one, are quite small, which makes the results promising, rather than conclusive.

Implications of the present study could be the following. The reported findings suggest that the applied procedures are useful supplements or alternatives in infant assessments. Additional research could be directed at the clinical practicality of the procedures and the adaptation of them to the requirements of different populations and individuals, that is, severely mentally and motorically or sensorily handicapped infants.

An extension of classical conditioning procedures could be found in the research of Rovee-Collier and collaborators and others (i.e., Gekoski, Fagen, & Pearlman, 1984; Little et al., 1984; Rovee-Collier & Fagen, 1981) and Moscovitch (1984), where retention of learned associations (in operant and classical conditioning paradigms) is suggested to be an important issue in infants' cognition. In fact, research with similar learning measures, populations and research questions points to the relation, and thus possible significance, of retention of a learned response (Little et al., 1984).

Another interesting field of research is suggested by Meltzoff (1985, 1988a, 1988b), who investigated on imitative processes in very young infants. Within the latter research topics the age of assessments is considered very critical, that is, these skills develop rather fast and ergo can be observed in a limited age range.

Fagen et al. (1987) conclude from their studies that infant memory is a relatively stable component of infant cognition that is related to later intelligence. Their results also suggest that measures of infant memory should be incorporated into any new attempts to develop screening devices for infants at risk for retarded mental development. It is reminded that by the investigation of the learning rate, at the same time (implicitly) also (short-term) memory is tested. Recent theories on memory emphasize the processes and strategies that lead to retention of presented information, that is, the way organisms learn (acquire knowledge), and retrieval of this acquired knowledge. Related to these issues are research efforts in infants directed at attribution of cause and effect (experiments by Zelazo, 1982; see also Lipsitt & Rovee-Collier, 1983, 1984). More in general, the contribution of cognitive information processing and the translation of it into new behavior (seen as adaptive learning processes) could be an area of research with classical and operant conditioning as the tools for studying these behaviors (cf. Davey, 1987).

An alternative source of information would be the investigation as to the concurrent validity of learning (conditioning) procedures by comparison of measures derived from those procedures with different assessment procedures applied at successive ages and different populations, especially in more homogeneous groups of handicapped infants (e.g., groups of Down's syndrome infants). The procedures could be used simultaneously or in parallel. Suggestions for these procedures could be (neuro)physiological procedures, like evoked potentials, also for gathering information on the development of learned responses and established infant tests. Also (conscious) attentive processes (studied through stimulus preference techniques) could be targets of study. Moreover, observations of the infants' interactions with their environment could possibly reveal relationships with the aforementioned measures.

In addition, different stimuli and responses, or more complex conditioning procedures, for example, discrimination conditioning paradigms, seem fruitful topics of investigation, not only for validation purposes, but also as assessment means for young multiply (sensory) handicapped children. Thus, these efforts could not only result in more reliable (cf. Fagan & McGrath, 1981) and valid infant measures, but also in an assessment battery that combines different techniques for measuring different aspects of infant development, including information processing (cf. Vietze, 1988).

Practically, the present results and the aforementioned considerations encourage continued research on the development of a more reliable test(battery) or screening device for assessment of present as well as subsequent cognitive level. Such a test would have a number of possible uses, of which screening of infants at risk and measurement of cognitive functioning in profoundly or severely retarded and/or sensorily handicapped populations for whom non-verbal tests are required, seem most prominent.

Abstract

Assessment of cognitive skills in very young children is considered important with respect to the estimate of present as well as future capacities. Especially relevant seems to be the aforementioned goal of assessment in high-risk and handicapped children. Generally one prefers to assess as early as possible for obvious reasons. Early assessment is mostly aimed at the measurement of present levels of functioning upon which eventual intervention can be based and at evaluation of the effects of such interventions, directed at preventing or overcoming developmental problems. In addition, such assessment could also contribute significantly to the design of (individual) intervention plans.

From the beginning of the century many efforts have been undertaken pursuing the aforementioned goals. All these efforts have resulted in standardized infant tests, analogous to those developed for children and adults, and more standardization in neurological and neurophysiological assessments. Although these tests and assessments have led to more accurate estimates of infant status, prediction of future developmental level showed to be rather disappointing. In addition, there was a heightening interest and need in reliable assessment of cognitive functions and processes playing a role in the developing organism.

The aforementioned considerations and others have evolved in a wide field of research on infant cognitive functioning (e.g., sensory processing and learning). The knowledge within this area increased vastly and, perhaps more importantly, the view of the infant as having a passive role in the first phase of development was increasingly abandoned. Now the infant is seen as an active participator interacting in a complicated environment.

The historical background of the aforementioned developments and some relevant literature describing them unlock the present report. That is, a brief history and background of infant testing appears in Chapter 1 of this paper. In Chapter 2 recent developments in the area of traditional assessment procedures, including neurological measurements and several procedures, old and new, are outlined.

Chapter 3 provides an overview of new assessment strategies, possible reasons for the increased interest in them, and problems they had (and still have) to overcome. More specifically, neurophysiological, environmental, and attention measurements and conditioned learning assessments are sketched in sections 3.2., 3.3., 3.4., and 3.5., respectively. In addition, some results of the aforementioned procedures, pertaining to assessment of developmental potential or prediction of future development, are discussed in these sections.

After the aforementiond survey a longitudinal experiment is presented in Chapter 4. Classical and operant conditioning strategies for the assessment of learning rate and home observation measures were used to predict later psychological development, measured by an operant conditioning procedure and the Bayley Scales of Infant Development. A small group of normal infants and a group of high-risk and handicapped infants were involved. The investigations took place when the subjects were 4, 6, and 18 months old. That is, at the age of 4 and 6 months the predictor measures and at 18 months the outcome measures were taken. The rationale, method, and procedures of this experiment are outlined in sections 4.1. and 4.2. Furthermore, in section 4.3 the results of the aforementioned experiment are presented, that is, the data are described for the two subgroups of subjects and comparisons are made between them. In addition, the data are analysed by means of simple correlational and multiple multivariate statistical procedures. Moreover, the groups of successful and unsuccessful high-risk subjects (successfulness is defined in relation to the conditioning procedures) are compared on their performances on the Bayley Mental and Motor Scales. Finally, all subjects are classified according to performance on the Bayley Mental Scale. That is, through discriminant analysis group membership is predicted using a selected set of predictive measures.

The results showed that on all predictor as well as outcome measures, not including the home observation measures, the mean scores of the high-risk subgroup were at a disadvantage in comparison with the normal group. The latter differences were never, except for the Bayley Mental Scale scores, significant. Simple correlational and multiple multivariate regression analyses indicate that the scores on two of a total of three of the classical and both the operant procedures (predictor measures) accounted for a substantial amount of variance in the outcome measures, that is, particularly the scores on the Bayley Scales. Furthermore, both operant predictor measures showed moderate significant correlations with the operant outcome measure. Only one of the home observation variables used in this study showed a significant correlation with outcome variables. A combination of selected predictor variables accounted for significant portions of the variance in the outcome variables. A larger amount of accounted variance was found than generally established with the single procedures mentioned in the literature survey (e.g., attention and environmental measures).

The classical procedures showed to be easily applicable, that is, only one conditioning failure (i.e., a subject which did not successfully condition within the maximum number of training trials allowed) was observed. The operant procedures, however, were much more vulnerable to failure. More particularly, almost 6% of all the conditioning occasions was unsuccessful. These unsuccessful subjects were kept in the study sample by the use of allotted scores. Furthermore, high-risk children who conditioned successfully within the maximum number of training trials allowed were compared with those who did not (conditioning failures). It was concluded, that it had indeed been meaningful, in view of the goal of this study, to maintain these unsuccessful subjects instead of to remove them from the study samples. In addition, classification of subjects on the basis of the three most powerful predictor variables (by means of a discriminant analysis procedure) resulted in a reasonably high percentage of correct group-predictions (more than 80 percent).

In conclusion, in section 4.4 the summary and conclusions as well as some implications of the present study are provided. In addition, several suggestions for future research are made.

Key words: high-risk infant, assessment, classical and operant conditioning, learning rate, prediction of development.

Samenvatting

Het vaststellen van cognitieve vaardigheden in zeer jonge kinderen is van belang met betrekking tot het inschatten van huidige zowel als latere mogelijkheden. Deze doelstelling is belangrijk voor kinderen die zich normaal ontwikkelen. Nog belangrijker lijkt deze voor kinderen met een verhoogd risico met betrekking tot een stagnerende ontwikkeling of voor kinderen met reeds gediagnostiseerde handicaps. Over het algemeen wil men zo vroeg mogelijk een diagnose stellen. Vroeg-diagnostiek is er meestal in eerste instantie op gericht het niveau van funktioneren op verschillende, in die ontwikkelingsperiode belangrijk geachte, gebieden vast te stellen. Daarop volgend kan eventueel interventie plaats vinden. Een interventie kan als doelstelling hebben het voorkómen of minimaliseren van ontwikkelingsstoornissen. Bovendien kunnen tests de effekten van dergelijke interventies trachten te meten of zelfs een bijdrage leveren aan het opstellen van interventieprogramma's.

Vanaf het begin van deze eeuw heeft men zich intensief beziggehouden met eerdergenoemde doelstelling. Dit heeft geresulteerd in gestandaardiseerde tests voor jonge kinderen, zoals die ook bestaan voor oudere kinderen en volwassenen. Daarnaast heeft er een ontwikkeling plaatsgevonden in de neurologische en neurofysiologische diagnostiek, met name met betrekking tot standaardisatie van gebruikte schalen. Hoewel al deze diagnostische instrumenten het mogelijk maakten het ontwikkelingsniveau van jonge kinderen nauwkeuriger vast te stellen, bleek de voorspellende waarde ervan ten aanzien van latere niveaus (of IQ) nog erg moeilijk. Bovendien ontstond er een groeiende behoefte aan en interesse voor het betrouwbaar nagaan van cognitieve funkties en processen, die een rol spelen in het zich ontwikkelende organisme.

De eerdergenoemde ontwikkelingen hebben tot gevolg gehad, dat er een breed onderzoeksterrein is ontstaan dat betrekking heeft op cognitieve processen in zeer jonge kinderen (van zuigelingen tot kleuters). De kennis op dit gebied nam snel toe en, wat nog belangrijker was, de opvatting dat een kind in zijn eerste levensfase met betrekking tot zijn eigen ontwikkeling een passieve rol speelde, veranderde. Nu ziet men het kind eerder als een zich aktief interacterend organisme in een complexe omgeving.

Het voorliggende rapport gaat in op de historische achtergrond van bovengenoemde ontwikkelingen en vermeldt enige relevante literatuur. In Hoofdstuk 1 wordt in kort bestek de geschiedenis van het testonderzoek van zeer jonge kinderen beschreven over de periode van het begin van deze eeuw tot in de zestiger jaren. Van een aantal belangrijke ontwikkelingsschalen, die in deze periode werden ontwikkeld, worden de belangrijkste kenmerken genoemd. In Hoofdstuk 2 worden recente ontwikkelingen op het gebied van traditioneel testonderzoek, waaronder neurologische onderzoek, geschetst. Verschillende testinstrumenten worden nader besproken.

Een aantal vanaf de zestiger jaren ontwikkelde procedures om zeer jonge kinderen te testen, wordt besproken in Hoofdstuk 3. Bovendien worden een aantal redenen genoemd voor de stijgende belangstelling in deze methodes. Ook worden een aantal problemen die aan deze methodes kleven genoemd. In de paragrafen 3.2., 3.3., 3.4., en 3.5. worden procedures die respektievelijk neurofysiologische-, omgevings- en attentievariabelen meten, en conditionerings-methoden behandeld. Onder andere worden resultaten besproken van de laatstgenoemde procedures, voor zover deze betrekking hebben op het vaststellen van individuele verschillen of het voorspellen van toekomstige ontwikkeling.

Na het overzicht van deze procedures wordt in Hoofdstuk 4 verslag gedaan van een longitudinaal onderzoek. Klassieke en operante conditionerings-strategieën zijn in dit onderzoek de belangrijkste bepalers van de leersnelheid van zeer jonge kinderen. De op deze manier verkregen scores worden gebruikt om de latere psychologische ontwikkeling (gemeten met behulp van een operante conditioneringsprocedure en de Bayley Ontwikkelingsschalen) te voorspellen. De subjecten zijn een kleine groep normale en een grotere groep zogenoemde risicokinderen. De onderzoeken vonden plaats toen de kinderen respektievelijk 4 en 6 maanden (voor de predictoren) en 18 maanden (voor de uitkomstvariabelen) oud waren. In de paragrafen 4.1. en 4.2. worden de methode en gehanteerde procedures van dit onderzoek beschreven. De resultaten van het onderzoek worden gepresenteerd in paragraaf 4.3. Eerst worden de data van de twee groepen vermeld en met elkaar vergeleken. Deze data worden geanalyseerd met behulp van correlationele en multipele multivariate statistische procedures. Vervolgens worden twee subgroepen van de risicogroep (nl. de succesvolle en niet-succesvolle kinderen, waarbij succes werd gedefinieerd in relatie tot de conditioneringsprocedures) met elkaar vergeleken op de scores op de Mentale Schaal van de Bayley. Tenslotte worden alle subjecten, op grond van een set geselecteerde predictoren, geklassificeerd met betrekking tot de scores op de Mentale Schaal van de Bayley met behulp van een discriminant analyse.

De resultaten lieten minder gunstige scores (hoewel voor de meeste variabelen niet statistisch significant) zien voor de risicokinderen. Uit regressieanalyses bleken substantiële proporties variantie in de uitkomstvariabelen verklaard te worden door verscheidene predictoren, met name wanneer een aantal predictorvariabelen gecombineerd werd.

De klassieke conditioneringsprocedures bleken makkelijk toe te passen, dat wil zeggen, er kwamen weinig gevallen voor waarin een subject niet geconditioneerd werd binnen een van te voren vastgesteld maximaal aantal trainingstrials (deze subjecten werden niet-succesvolle conditioneerders genoemd). De operante procedures echter bleken kwetsbaarder ten aanzien van dit laatste criterium. De subjecten, die niet aan dit criterium voldeden, kregen scores toegewezen, zodat ze niet uitgesloten waren van de data-analyses. Bovendien werden de risicokinderen geklassificeerd op basis van de sterkste predictoren met behulp van discriminantanalyses. Deze analyses lieten een vrij hoog percentage juiste predikties zien (meer dan 80 percent).

Tenslotte worden in paragraaf 4.4. een korte samenvatting, een bespreking van de resultaten en enige implicaties van dit onderzoek gegeven. Bovendien worden enige suggesties voor verder onderzoek gedaan.

Trefwoorden: risicokinderen, diagnostiek, klassieke en operante conditionering, leersnelheid, voorspelling van ontwikkeling.

References

Allen, M.J., & Yen, W.M. (1979). *Introduction to measurement theory.* Monterey, CA: Brooks/Cole.

Als, H., & Duffy, F.H. (1983). The behavior of the premature infant: a theoretical framework for a systematic assessment. In T.B. Brazelton & B.M. Lester (Eds.), *New approaches to developmental screening of infants* (pp. 153-173). New York: Elsevier.

Als, H., Lester, B.M., Tronick, E.Z., & Brazelton, B.M. (1982). Toward a research instrument for the assessment of preterm infants' behavior (A.P.I.B.). In H.E. Fitzgerald, B.M. Lester & M.W. Yogman (Eds.), *Theory and research in behavioral pediatrics: Vol. 1* (35-132). New York: Plenum.

Bakeman, R., & Brown, J.V. (1980). Early interaction: consequences for social and mental development at three years. *Child Development, 51,* 437-447.

Barnard, K., Bee, H.L., & Hammond, M.A. Home environment and cognitive development in a healthy, low-risk sample: The Seattle study (1984). In A.W. Gottfried (Ed.), *Home environment and early cognitive development* (pp. 117-149). New York: Academic.

Barrera, M.E., Rosenbaum, P.L., & Cunningham, C.E. (1986). Early home intervention with low-birth-weight infants and their parents. *Child Development, 57,* 20-33.

Barrera, M.E., Rosenbaum, P.L., & Cunningham, C.E. (1987). Corrected and uncorrected Bayley scores: Longitudinal developmental patterns in low and high birthweight preterm infants. *Infant Behavior and Development, 10,* 337-346.

Bathurst, K., & Gottfried, A.W. (1987). Untestable subjects in child development research: developmental implications. *Child Development, 58,* 1135-1144.

Bayley, N. (1967). The growth of intelligence. In Y. Brackbill & G.G. Thompson (Eds.), *Behavior in infancy and childhood. A book of readings* (pp. 398-417). New York: The Free Press.

Bayley, N. (1969). *Manual for the Bayley Scales of Infant Development.* New York: The Psychological Corporation.

Beckwith, L., & Cohen, S.E. (1984). Home environment and cognitive competence in preterm children during the first 5 years. In A.W. Gottfried (Ed.), *Home environment and early cognitive development* (pp. 235-271). New York: Academic.

Beckwith, L., & Parmelee, A.H. (1986). EEG patterns of preterm infants, home environment, and later IQ. *Child Development, 57,* 777-789.

Bee, H.L., Barnard, K.E., Eyres, S.J., Gray, C.A., Hammond, M.A., Spietz, A.L., Snyder, C., & Clark, B. (1982). Prediction of IQ and language skill from perinatal status, child performance, family characteristics, and mother-infant interaction. *Child Development*, *53*, 1134-1156.

Berg, W.K., & Berg, K.M. (1979) Psychophysiological development in infancy: state, sensory function, and attention. In J. Osofsky (Ed.), *Handbook of infant development* (pp. 283-343). New York: Wiley.

Bernard, P.A. (1985). The efficacy of brainstem response audiometry in the diagnosis of meningitis and other CNS pathology. In S.E. Trehub & B. Schneider (Eds.) *Auditory development in infancy* (pp. 157 164). New York: Plenum.

Bierman-van Eendenburg, M.E.C., Jurgens-van der Zee, A.D., Olinga, A.A., Huisjes, H.H., & Touwen, B.C.L. (1981). Predictive value of neonatal neurological examination: a follow-up study at 18 months. *Developmental Medicine & Child Neurology*, *23*, 296-305.

Blalock, H.M. (1979). *Social statistics*. Tokyo: McGraw-Hill.

Blass, E.M., Ganchrow, J.R., & Steiner, J.E. (1984). Classical conditioning in newborn humans 2-48 hours of age. *Infant Behavior and Development*, *7*, 223-235.

Block, J.D., Sersen, E.A., & Wortis, J. (1970). Cardiac classical conditioning and reversal in the mongoloid, encephalopathic, and normal child. *Child Development*, *41*, 771-785.

Bornstein. M.H., & Benasich, A.A. (1986). Infant habituation: assessments of individual differences and short-term reliability at five months. *Child Development*, *57*, 87-99.

Bower, G.H., & Hilgard, E.R. (1981). *Theories of learning*. London: Prentice Hall.

Bower, T.G.R. (1989). *The rational infant*. New York: Freeman.

Brackbill, Y., Fitzgerald, H.E., & Lintz, L.M. (1967). A developmental study of classical conditioning. *Monographs of the Society for Research in Child Development*, serial no. 116, vol. 32, no. 8.

Brackbill, Y., & Fitzgerald, H.E. (1972). Stereotype temporal conditioning in infants. *Psychophysiology*, *9*, 569-577.

Bradley, R.H., & Caldwell, B.M. (1976). Early home environment and changes in mental test performance in children from 6 to 36 months. *Developmental Psychology*, *12*, 93-97.

Bradley, R.H., & Caldwell, B.M. (1978). Screening the environment. *American Journal of Orthopsychiatry*, *48*, 114-130.

Bradley, R.H., & Caldwell, B.M. (1980). The relation of home environment, cognitive competence, and IQ among males and females. *Child Development*, *51*, 1140-1148.

Bradley, R.H., & Caldwell, B.M. (1982). The consistency of the home environment and its relation to child development. *International Journal of Behavioral Development*, *5*, 445-465.

Bradley, R.H., & Caldwell, B.M. (1984a). The relation of infants' home environments to achievement test performance in first grade: a follow-up study. *Child Development*, *55*, 803-809.

Bradley, R.H., & Caldwell, B.R. (1984b). 174 Children: a study of the relationship between home environment and cognitive development during the first 5 years. In A. W. Gottfried (Ed.), *Home environment and early cognitive development* (pp. 5-56). New York: Academic.

Bradley, R.H., & Caldwell, B.R., & Rock, S.L. (1988). Home environment and school performance: A ten year follow-up and examination of three models of environmental action. *Child Development, 59*, 852-867.

Bradley, R.H., & Caldwell, B.R., Rock, S.L., Barnard, K. E., Gray, C., Hammond, M.A., Mitchell, S., Siegel, L., Ramey, C. T., Gottfried, A.W., & Johnson, D.L. (1989). Home environment and cognitive development in the first three years of life: A collaborative study involving six sites and three ethnic groups in North America. *Developmental Psychology, 25*, 217-235.

Bradley-Johnson, S., & Travers, M.W. (1979). Cardiac change of retarded and nonretarded infants to an auditory signal. *American Journal of Mental Deficiency, 83(6)*, 631-636.

Brazelton, T.B. (1973). *Neonatal behavioral assessment scale*. Clinics in Developmental Medicine no. 50. London: Spastics International Medical Publications.

Bricker, D.D. (Ed.). (1982). *Intervention with at risk and handicapped infants*. Baltimore: University Park.

Brooks, J., & Weinraub, M. (1976). A history of infant intelligence testing. In M. Lewis (Ed.), *Origins of intelligence* (pp. 19-58). New York: Wiley.

Carlson, L., & Bricker, D.D. (1982). Dyadic and contingent aspects of early communicative intervention. In D.D. Bricker (Ed.), *Intervention with at risk and handicapped infants* (pp. 291-308). Baltimore: University Park.

Caron, A.J., Caron, R.F., & Glass, P. (1983). Responsiveness to relational information as a measure of cognitive functioning in nonsuspect infants. In T. Field & A. Sostek (Eds.) *Infants born at risk. Physiological, perceptual and cognitive processes* (pp. 181-209). New York: Grune & Stratton.

Cattell, P. (1966). *The measurement of intelligence of infants and young children*. New York: The Psychological Corporation.

Clifton, R.K. (1974). Heart rate conditioning in the newborn infant. *Journal of Experimental Child Psychology, 18*, 9-21.

Coates, D.L., & Lewis, M. (1984). Early mother-infant interaction and infant cognitive status as predictors of school performance and cognitive behavior in six-year olds. *Child Development, 55*, 1219-1230.

Cohen, L.B. (1981). Examination of habituation as a measure of aberrant infant development. In S.L. Friedman & M. Sigman (Eds.), *Preterm birth and psychological development* (241-253). New York: Academic.

Cohen, S.E., & Beckwith, L. (1979). Preterm infant interaction with the caregiver in the first year of life and competence at age two. *Child Development, 50*, 767-776.

Cohen, S.E., & Parmelee, A.H. (1983). Prediction of five-year Stanford-Binet Scores in preterm infants. *Child Development, 54*, 1242-1253.

Connolly, K., & Stratton, P. (1969). An exploration of some parameters affecting classical conditioning in the neonate. *Child Development, 40*, 431-441.

Cools, A.T.M., & Hermanns, J.M.A. (1977). *Vroegtijdige onderkenning van problemen in de ontwikkeling van kinderen*. Amsterdam: Swets & Zeitlinger.

Costello, A.M. de L., Hamilton, P.A., Baudin, J., Townsend, J., Bradford, B.., Stewart, A.L., & Reynolds, E.O.R. (1988). Prediction of neurodevelopmental impairment at four years from brain ultrasound appearance of very preterm infants. *Developmental Medicine and Child Neurology, 30*, 711-722.

Crowell, D.H., Blurton, L.B., Kobayashi, J.R., McFarland, J.Z., & Yang, R.K. (1976). Studies in early infant learning: Classical conditioning of the neonatal heart rate. *Developmental Psychology Monograph, 12*, 373-397.

Davey, G. (1981). *Animal learning and conditioning*. London: McMillan.

Davey, G. (Ed.). (1987). *Cognitive processes and Pavlovian conditioning in humans*. Chichester: Wiley.

Davis, A. (1985). Electric response audiometry in young children. In S.E. Trehub & B. Schneider (Eds.), *Auditory development in infancy* (pp. 165-175). New York: Plenum.

Dawson, M.E., & Schell, A.M. (1987). Human autonomic and skeletal classical conditioning: The role of conscious cognitive factors. In G. Davey (Ed.), *Cognitive processes and Pavlovian conditioning in humans* (pp.27-55). Chichester: Wiley.

Desmedt, J.E. (Ed.). (1977). *Visual evoked potential in man: New developments*. Oxford: Clarendon.

Despland, P.A., & Galambos, R. (1980a). The Auditory Brainstem Response (ABR) is a useful tool in the intensive care nursery.
Pediatric Research, 14, 154-158.

Despland, P.A., & Galambos, R. (1980b). Use of the auditory brainstem responses by prematures and newborn infants. *Neuropediatrics, 11*, 99-107.

Dreyfus-Brisac, C., & Ellingson, R.J. (1977a). Prognostic value of the neonatal EEG in full-term newborns. In D.D. Daly (Ed.), *Clinical EEG, V.: Vol. 15B* (pp. 89-100). In A. Remond (Ed.), *Handbook of electroencephalography and clinical neurophysiology*. Amsterdam: Elsevier.

Dreyfus-Brisac, C., & Ellingson, R.J. (1977b). Sensory evoked responses in the neonatal period and their application. In D.D. Daly (Ed.), *Clinical EEG, V.: Vol. 15B* (pp. 74-88). In A. Remond (Ed.), *Handbook of electroencephalography and clinical neurophysiology*. Amsterdam: Elsevier.

Drillien, G.M. (1972). Aetiology and outcome in low-birthweight infants. *Developmental Medicine and Child Neurology, 14*, 563-574.

Drotar, D., & Sturm, L. (1988). Prediction of intellectual development in young children with early histories of nonorganic failure-to-thrive. *Journal of Pediatric Psychology*, 2, 281-296.

Dubowitz, L.M.S., Dubowitz, V., & Goldberg, C. (1970). Clinical assessment of gestational age in the newborn infant. *Journal of Pediatrics, 77*, 1-10.

Dubowitz, L.M.S., Dubowitz, V., Palmer, P.G., Miller, G., Fawer, C.-L., & Levene, M.I. (1984). Correlation of neurologic assessment in the preterm newborn infant with outcome at 1 year. *The Journal of Pediatrics, 105*, 452-456.

Duffy, F.H., & Als, H. (1983). Neurophysiological assessment of the neonate: an approach combining Brain Electrical Activity Mapping (BEAM) with Behavioral

Assessment (APIB). In T.B. Brazelton & B.M. Lester (Eds.), *New approaches to developmental screening of infants* (pp. 175-196). New York: Elsevier.
Eggermont, J.J. (1985). To BER or not to BER. In S.E. Trehub & B. Schneider (Eds.), *Auditory development in infancy* (pp. 177-180). New York: Plenum.
Eldredge, L., & Salamy, A. (1988). Neurobehavioral and neurophysiological assessment of healthy and "at-risk" full-term infants. *Child Development*, *59*, 186-192.
Ellingson, R.J., Dutch, S.J., & McIntire, M.S. (1974). EEG's of prematures: 3-8 year follow-up study. *Developmental Psycho-biology*, *7(6)*, 529-538.
Evers-Embden, B., & Scholte, E.M. (1983). Leerprestaties en neurologische status praesens van prematuren, premature-dysmaturen en dysmaturen op 8-jarige leeftijd. *Tijdschrift Kindergeneeskunde*, *51*, 85-94.
Fagan, J.F. III (1984). Infant memory. History, current trends, relations to cognitive psychology. In M. Moscovitch (Ed.), *Infant memory* (pp. 1-27). New York: Plenum.
Fagan, J.F. III (1988). Screening infants for later mental retardation: From theory to practice. In P.M. Vietze & H.G. Vaughan (Eds.), *Early identification of infants with developmental disabilities* (pp.253-265). Philadelphia: Grune & Stratton.
Fagan, J.F. III, & McGrath, S.K. (1981). Infant recognition memory and later intelligence. *Intelligence*, *5*, 121-130.
Fagan, J.F. III, & Singer, L.T. (1983). Infant recognition memory as a measure of intelligence. In L.P. Lipsitt (Ed.), *Advances in infancy research: Vol. 2* (pp. 31-78). Norwood: Ablex.
Fagen, J.W., Ohr, P.S., Fleckenstein, L.K., & Singer, J.M. (1987). *Infant long-term memory for a conditioned response and intelligence test performance at 2 years of age*. Paper presented at the Society for Research in Child Development, Baltimore.
Field, T.M., Hallock, N., Ting, G., Dempsey, J., Dabiri, C., & Shuman, H.H. (1978). A first-year follow-up of high-risk infants: Formulating a cumulative risk index. *Child Development*, *49*, 119-131.
Field, T.M., Huston, A., Quay, H.C., Troll, L., & Finley, G.E. (Eds.).(1982). *Review of human development*. New York: Wiley.
Field, T.M., Sostek, A.M., Goldberg, S., & Shuman, M.D. (Eds.). (1979). *Infants born at risk. Behavior and development*. New York: Spectrum Publications Medical & Scientific Books.
Finkelstein, N.W., Gallagher, J.J., & Farran, D.C. (1980). Attentiveness and responsiveness to auditory stimuli of children at risk for mental retardation. *American Journal of Mental Deficiency*, *85*, 135-144.
Finkelstein, N.W., & Ramey, C.T. (1977). Learning to control the environment in infancy. *Child Development*, *48*, 806-819.
Fitzgerald, H.E., & Brackbill, Y. (1976). Classical conditioning in infancy: Development and constraints. *Psychological Bulletin*, *83(3)*, 353-376.
Fitzgerald, H.E., Lintz, L.M., Brackbill, Y., & Adams, G. (1967). Time perception and conditioning an autonomic response in human infants. *Perceptual and Motor Skills*, *24*, 479-486.
Fitzgerald, H.E., & Porges, S.W. (1971). A decade of infant conditioning and learning research. *Merrill-Palmer Quarterly*, *17*, 79-117.

Fitzhardinge, P.M., Flodmark, O., Fitz, C.R., & Ashby, S. (1981). The prognostic value of computed tomography as an adjunct to assessment of the term infant with postasphyxial encephalopathy. *The Journal of Pediatrics, 99 (5)*, 777-781.

Flavell, J.H. (1985). *Cognitive development* (2nd ed.). Englewood Cliffs: Prentice Hall.

Francis-Williams, J. (1977). Psychological assessment. In C.M. Drillien & M.B. Drummond (Eds.), *Neurodevelopmental problems in early childhood. Assessment and management* (pp. 110-125). Oxford: Blackwell Scientific Publications.

Frankenburg, W.K., & Coons, C.E. (1983). Early identification of at risk children. In T.B. Brazelton & B.M. Lester (Eds.), *New approaches to developmental screening of infants* (pp. 137-152). New York: Elsevier.

Frankenburg, W.K., & Dodds, J.B. (1967). The Denver Developmental Screening Test. *The Journal of Pediatrics*, *71*, 181-191.

Franks, V., & Franks, C.M. (1962). Classical conditioning procedures as an index of vocational adjustment among mental defectives. *Perceptual and Motor Skills*, *14*, 241-242.

Friedman, M.P., Das, J.P., & O'Connor, N. (1981). *Intelligence and learning*. New York: Plenum.

Friedman, S.L., Jacobs, B.S., & Werthmann, M.W.. (1981). Sensory processing in pre- and full-term infants in the neonatal period. In S.L. Friedman & M. Sigman (Eds.), *Preterm birth and psychological development* (pp. 159-178). New York: Academic.

Gambi, D., Rossini, P.M., Albertini, G., Sollazzo, D., Torrioli, M.G., & Polidori, G.C. (1980). Follow-up of visual evoked potential in full-term and preterm control newborns and in subjects who suffered from perinatal respiratory distress. *Electroencephalography and Clinical Neurophysiology*, *48*, 509-516.

Gardner, J.M., & Karmel, B.Z. (1983). Attention and arousal in preterm and full-term neonates. In T. Field & A. Sostek (Eds.), *Infants born at risk: Psychological and perceptual processes* (pp. 69-98). New York: Grune & Stratton.

Gaussen, T. (1984). Developmental milestones or conceptual milestones? Some practical and theoretical limitations in infant assessment procedures. *Child: Care, Health and Development*, *10*, 99-115.

Gaussen, T., & Stratton, P. (1985). Beyond the milestone model - a systems framework for alternative infant assessment procedures. *Child: Care, Health and Development*, *11*, 131-150.

Gekoski, M.J., & Fagen, J.W. (1984). Noncontingent stimulation, stimulus familiarization and subsequent learning in young infants. *Child Development*, *55*, 2226-2233.

Gekoski, M.J., Fagen, J.W., & Pearlman, M.A. (1984). Early learning and memory in the preterm infant. *Infant Behavior and Development*, *7*, 267-276.

Gesell, A., & Amatruda, C.S. (1947). *Developmental Diagnosis*. New York: Hoeber.

Gewirtz, J. L., & Petrovich, S.B. (1982). Early social and attachment learning in the frame of organic and cultural evolution. In T. Field, A. Huston, A. Quay, L. Troll & G. Finley (Eds.), *Review of human development* (pp. 3-19). New York: Wiley.

Gold, P. (1979). Suspected neurological impairment and cognitive abilities: a longitudinal study. *Psychological Reports*, *45*, 215-218.

Goldberg, S., & Kearsley, R.B. (1983). Guest editorial. *Child Development, 54*, 1083-1085.

Golden, M., & Birns, B. (1976). Social class and infant intelligence. In M. Lewis (Ed.), *Origins of intelligence* (pp. 299-351). New York: Wiley.

Gollin, E.S. (1981). *Developmental Plasticity. Behavioral and biological aspects of variations in development*. New York: Academic.

Gottlieb, G., & Krasnegor, N.A. (Eds.).(1985). *Measurement of audition and vision in the first year of postnatal life. A methodological overview*. Norwood: Ablex.

Graham, F.K., Matarazzo, R.G., & Caldwell, B.M. (1956). Behavioral differences between normal and traumatized newborns:II. Standardization, reliability, and validity. *Psychological Monographs, 70*, (21, Whole No. 428).

Greenberg, M.T., & Crnic, K.A. (1988). Longitudinal predictors of developmental status and social interaction in premature and full-term infants at age two. *Child Development, 59*, 554-570.

Griffiths, R. (1954). *The abilities of babies*. London: University of London.

Griffiths, R. (1970). *The abilities of young children*. London: Child Development Research Center.

Grings, W.W., Lockhart, R.A., & Dameron, L.E. (1962). Conditioning autonomic responses of mentally subnormal individuals. *Psychological Monographs: General and Applied, 76*, (39, Whole no. 558).

Hadders-Algra, M., Huisjes, H.J., & Touwen, B.C.L. (1988a). Perinatal correlates of major and minor neurological dysfunction at school age: A multivariate analysis. *Developmental Medicine and Child Neurology, 30*, 472-481.

Hadders-Algra, M., Huisjes, H.J., & Touwen, B.C.L. (1988b). Perinatal risk factors and minor neurological dysfunction: Significance for behaviour and school achievement at nine years. *Developmental Medicine and Child Neurology, 30*, 482-491.

Hall, J.F. (1982). *An invitation to learning and memory*. Boston: Allyn and Bacon.

Hamers, J.H.M., & Ruijssenaars, A.J.J.M. (1984). *Leergeschiktheid en leertests*. Harlingen: Flevodruk.

Hartmann, D. P. (1984). Assessment strategies. In D.H. Barlow & M. Hersen, *Single case experimental designs* (2nd ed) (pp. 107-139). New York: Pergamon.

Hayes, L.A., Ewy, R.D., & Watson, J.S. (1982). Attention as a predictor of learning in infants. *Journal of Experimental Child Psychology, 34*, 38-45.

Hecox, K.E., Cone, B., & Blaw, M.E. (1981). Brainstem auditory evoked response in the diagnosis of pediatric neurologic diseases. *Neurology, 31*, 832-840.

Hecox, K.E., & Deegan, D.M. (1985). Methodological issues in the study of auditory development. In G. Gottlieb & N.A. Krasnegor (Eds.), *Measurement of audition and vision in the first year of postnatal life* (pp. 367-389). Norwood: Ablex.

Henderson, N.B., & Engel, R. (1974). Neonatal visual evoked potentials as predictor of psychoeducational tests at age seven. *Developmental Psychology, 10*, 269-276.

Hoffman, H.S., Cohen, M.E., & DeVido, C.J. (1985). A comparison of classical eyelid conditioning in adults and infants. *Infant Behavior and Development, 8*, 247-254.

Honzik, M.P. (1976). Value and limitations of infant tests: an overview. In M. Lewis (Ed.), *Origins of intelligence* (pp. 59-95). New York: Wiley.

Horowitz, F.D., & Paden, L.Y. (1975). The effectiveness of environmental intervention programs. In B.M. Caldwell and Ricutti (Eds.), *Review of child development research*. Chicago: University of Chicago.

Horwitz, S.J., & Amiel-Tison, C. (1979). Neurologic problems. In M. H. Klaus and A. A. Fanaroff (Eds.), *Care of the high-risk infant* (pp. 360-381). Philadelphia: Saunders.

Hrbek, A., Karlberg, P., Kjellmer, I., Olsson, T., & Rika, M. (1977). Clinical application of evoked encephalographic responses in newborn infants. I: Perinatal asphyxia. *Developmental Medicine and Child Neurology*, *19*, 34-44.

Hunt, J.V. (1979a). Longitudinal research: A method for studying the intellectual development of high-risk preterm infants. In T.M. Field, A.M. Sostek, S. Goldberg & H.H. Shuman (Eds.), *Infants born at risk* (pp. 443-459). New York: SP Medical & Scientific Books.

Hunt, J.McV. (1979b). Psychological development: early experience. *Annual Review of Psychology*, *30*, 103-143.

Hunt, J. V. (1981). Predicting intellectual disorders in childhood for preterm infants with birthweights below 1501 grams. In S.L. Friedman & M. Sigman (Eds.), *Preterm birth and psychological development* (pp. 329-351). New York: Academic.

Illingworth, R.S. (1975). *The development of the infant and young child*. Edinburgh: Churchill Livingstone.

Illingworth, R.S. (1983). *The normal child*. Edinburgh: Churchill Livingstone.

Ingram, E., & Fitzgerald, H.E. (1974). Individual differences in infant orienting and autonomic conditioning. *Developmental Psychobiology*, *7*, 359-367.

Irzhanskaia, K.N., & Felberbaum, R.A. (1967). Effects of stimulus sensitization on ease of conditioning. In Y. Brackbill & G.G. Thompson (Eds.), *Behavior in infancy and early childhood* (pp. 446-249). New York: The Free Press.

Jackson, E. (1982). Environments of high-risk and handicapped infants. In C.T. Ramey & P.L. Trohanis (Eds.), *Finding and educating high-risk and handicapped infants* (pp. 53-67). Baltimore: University Park.

Jaffe, B. F. (Ed.). (1976). *Hearing loss in children*. Baltimore: University Park.

Jay, S., & Farran, D.C. (1981). The relative efficacy of predicting IQ from mother-child interactions using ratings versus behavioral count measures. *Journal of Applied Developmental Psychology*, *2*, 165-177.

Jensen, A.R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, *39*, 1.

Kail, R. (1984) *The development of memory and cognition in children* (2nd ed.). New York: Freeman.

Karch, D., Rothe, R., Jurisch, R., Heldt-Hildebrand, R., Luebbesmeier, A., & Lemburg, P. (1982). Behavioural changes and bioelectric brain maturation of preterm and fullterm newborn infants: a polygraph study. *Developmental Medicine and Child Neurology*, *24*, 30-47.

Karmel, B.Z., Gardner, J.M., Brown, E.G., Zappulla, R.A., Magnero, C.A., Goodman, J.A., & Carr, L.Y. (1988). In P.M. Vietze & H.G. Vaughan (Eds.), *Early identification of infants with developmental disabilities* (pp. 181-209). Philadelphia: Grune & Stratton.

Kirk, R.E. (1982) *Experimental design: Procedures for the behavioral sciences* (2nd ed.). Monterey (CA): Brooks/Cole.

Klein, P.S. (1984). Behavior of Israeli mothers toward infants in relation to infants' perceived temperament. *Child Development*, *55*, 1212-1218.

Knobloch, H., & Pasamanick, B. (Eds.). (1974). *Gesell and Amatruda's Developmental Diagnosis*. Hagerstown: Harper & Row.

Knobloch, H., Stevens, F., & Malone, A.F. (1980). *Manual of developmental diagnosis. The administration and interpretation of the revised Gesell and Amatruda. Developmental & neurologic examination*. Hagerstown: Harper & Row.

Kopp, C.B., & McCall, R.B. (1982). Predicting later mental performance for normal, at-risk and handicapped infants. In P.B. Baltes & O.G. Brim (Eds.), *Life-span development and behavior: Vol 4* (pp. 33-61). New York: Academic.

Kopp, C.B., & Vaughn, B.E. (1982). Sustained attention during exploratory manipulation as a predictor of cognitive competence in preterm infants. *Child Development*, *53*, 174-182.

Kotlarek, F., Zeumer, H., & Hornchen, A. (1980). Computertomographische Untersuchungen bei Enzephalopatien in der Neugeborenenperiode: Ausgangsbefunde und Verlaufsbeobachtungen. *Neuropaediatrie*, *11(2)*, 121-138.

Kurzberg, D., Stapells, D.R., & Wallace, I.F. (1988). Event-related potential assessment of auditory system integrity: Implications for language development. In P.M. Vietze & H.G. Vaughan (Eds.), *Early identification of infants with developmental disabilities* (pp.160-180). Philadelphia: Grune & Stratton.

Lancioni, G.E. (1978). *Conditioning in infancy and implications for intervention programs*. Lawrence, KS; Kansas Research Institute.

Lancioni, G.E. (1980). Infant operant conditioning and its implications for early intervention. *Psychological Bulletin*, *88*, 516-534.

Lancioni, G.E., Coninx, F., & Smeets, P. (1989). A classical conditioning procedure for the hearing assessment of multiple handicapped persons. *Journal of Speech and Hearing Disorders*, *54*, 88-93.

Lancioni, G.E., & Hoogland, G.A. (1980). Hearing assessment in young infants by means of a classical conditioning procedure. *International Journal of Pediatric Otorhinolaryngology*, *2*, 193-200.

Lancioni, G.E., Horowitz, F.D., & Sullivan, J. (1980). NBAS-K:I. A study of its stability and structure over the first month of life. *Infant Behavior and Development*, *3*, 341-359.

Lancioni, G.E., Hoogland, G.A., Smeets, P.M., Brozzi, G. Scoponi, M.V., Piatella, L., & Zamponi, N. (1985). Hearing assessment in developmentally impaired infants: classical conditioning as a supplement to brainstem-evoked response audiometry (BERA). *International Journal of Pediatric Otorhinolaryngology*, *10*, 221-228.

Lester, M.L., Karmel, B.Z., Cantor, D.S., & Wheeler, M.B. (1983). Application of pattern evoked potential techniques for evaluating infant perceptual systems. *Journal of Clinical Neuropsychology*, *5*, 39-50.

Lewis, M., & Brooks-Gunn, J. (1981). Visual attention at three months as a predictor of cognitive functioning at two years of age. *Intelligence*, *5*, 131-140.

Lewis, M., & Taft, L.T. (Eds.). (1982). *Developmental disabilities. Theory, assessment and intervention*. Lancaster: MTP.
Lintz, L.M., Fitzgerald, H.E., & Brackbill, Y. (1967). Conditioning the eyeblink response to sound in infants. *Psychonomic Science, 7*, 405-406.
Lipsitt, L.P. (1969). Learning capacities of the human infant. In R.J. Robinson (Ed.), *Brain and early behavior* (pp.227-249). London: Academic.
Lipsitt, L.P. (1979). Learning assessments and interventions for the infant born at risk. In T.M. Field, A.M. Sostek, S. Goldberg, & H.H. Shuman (Eds.), *Infants born at risk* (pp. 145-169). New York: SP Medical & Scientific Books.
Lipsitt, L.P. (1982). Infant learning. In T.M. Field, A. Huston, A.C. Quay, L. Troll, & G.E. Finley (Eds.), *Review of human development* (pp. 62-78). New York: Wiley.
Lipsitt, L.P., & Rovee-Collier, C.K. (Eds.). (1983). *Advances in infancy research: Vol. 2.* Norwood: N.J. Ablex.
Lipsitt, L.P., & Rovee-Collier, C.K. (Eds.). (1984). *Advances in infancy research: Vol. 3.* Norwood: N.J. Ablex.
Lipton, M.A. (1976). Early experience and plasticity in the CNS. In T.D. Tjossem (Ed.), *Intervention strategies for high risk infants and young children* (pp. 63-73). Baltimore: University Park.
Little, A.H., Lipsitt, L.P., & Rovee-Colier, C. (1984). Classical conditioning and retention of the infant's eyelid response: Effects of age and interstimulus interval. *Journal of Experimental Child Psychology, 37*, 512-524.
Lombroso, C.T., & Matsumiya, Y. (1985). Stability in waking-sleep states in neonates as a predictor of long-term neurologic outcome. *Pediatrics, 76*, 52-63.
Lubchenco, L.O. (1983). Intra-uterine growth assessment as predictor of neonatal risk. In T.B. Brazelton & B.M. Lester (Eds.), *New approaches for developmental screening of infants* (pp. 115-123). New York: Elsevier.
Majnemer, A., Rosenblatt, B., & Riley, P. (1988). Prognostic significance of the auditory brainstem evoked response in high-risk neonates. *Developmental Medicine and Child Neurology, 30*, 43-52.
Maltzmann, I., & Mandell, M.P. (1968). The orienting reflex as a predictor of learning and performance. *Journal of Experimental Research in Personality, 3*, 99-106.
Marquis, D.P. (1931). Can conditioned responses be established in the newborn infant? *Journal of Genetic Psychology, 39*, 479-492.
Marton, P., Minde, K., & Ogilvie, J. (1981). Mother-infant interactions in the premature nursery: a sequential analysis. In S. L. Friedman & M Sigman (Eds.), *Preterm birth and psychological development* (pp. 179-205). New York: Academic.
McCall, R.B. (1979). The development of intellectual functioning in infancy and the prediction of later IQ. In J.D. Osofsky (Ed.), *Handbook of infant development* (pp. 707-741). New York: Wiley.
McCall, R.B. (1981). Early predictors of later IQ: The search continues. *Intelligence, 5*, 141-147.
McCall, R.B. (1982). Issues in the early development of intelligence and its assessment. In M. Lewis & L.T. Taft (Eds.), *Developmental disabilities. Theory, assessment and intervention* (pp. 177-184). Lancaster: MTP.

McCall, R.B. (1983). Predicting developmental outcome. In T.B. Brazelton & B.M. Lester (Eds.), *New approaches for developmental screening of infants* (pp. 13-26). New York: Elsevier.

McCall, R.B., Hogarty, P.S., & Hurlburt, N. (1972). Transitions in infant sensorimotor development and the prediction of childhood IQ. *American Psychologist, 72*, 728-748.

Meltzoff, A.N. (1985). Immediate and deferred imitation in fourteen- and twenty-four-month-old infants. *Child Development, 56*, 62-72.

Meltzoff, A.N. (1988a). Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology, 24*, 470-476.

Meltzoff, A.N. (1988b). Infant imitation and memory: Nine-month-olds in immediate and deferred tasks. *Child Development, 59*, 217-225.

Meltzoff, A.N., & Moore, M.K. (1983). The origins of imitation in infancy: Paradigm, phenomena, and theories. In L.P. Lipsitt (Ed.), *Advances in infancy research: Vol. 2*. Norwood: Ablex.

Meulen, B.F. van der, & Smrkovsky, M. (1983). *BOS 2-30. Bayley Ontwikkelingsschalen. Handleiding*. Lisse: Swets & Zeitlinger.

Millar, S.W. (1976). Social reinforcement of a manipulative response in six- and nine-month-old infants. *Journal of Child Psychology and Psychiatry, 17*, 205-212.

Miller, D.J., Ryan, E.B., Aberger, E., McGuire, M.D., Short, E.J., & Kenny, D.A. (1979). Relationships between assessments of habituation and cognitive performance in the early years of life. *International Journal of Behavioral Development, 2*, 159-170.

Miller, D.J., Spiridigliozzi, G., Ryan, E.B., Callan, M.P., & McLaughlin, J.E. (1980). Habituation and cognitive performance: Relationships between measures at four years of age and earlier assessments. *International Journal of Behavioral Development, 3*, 131-146.

Miranda, S.B. (1976). Visual attention in defective and high-risk infants. *Merrill-Palmer Quarterly, 22*, 201-228.

Miranda, S.B., & Hack, M. (1979). The predictive value of neonatal visual-perceptual behaviors. In T.M. Field (Ed.), *Infants born at risk. Behavior and development* (pp. 69-90). New York: Spectrum Publications Medical & Scientific Books.

Mizrahi, E.M., & Dorfman, L.J. (1980). Sensory evoked potentials: clinical applications in pediatrics. *The Journal of Pediatrics, 97(1)*, 1-10.

Molfese, V.J., & Thomson, B. (1985). Optimality versus complications: Assessing predictive values of perinatal scales. *Child Development, 56*, 810-823.

Moscovitch, M. (Ed.).(1984). *Infant memory*. New York: Plenum.

Moss, M., Colombo, J., Mitchell, D.W., & Horowitz, F.D. (1988). Neonatal behavioral organization and visual processing at three months. *Child Development, 59*, 1211-1220.

Murakawi, R., Nakamura, H., Mizojiri, T.,Aida, M., & Matsuo, T. (1981). A study of brain development in low-birth-weight infants by computerized tomography. *Neuropediatrics, 12(2)*, 132-142.

Murray, A.D. (1988a). Newborn auditory brainstem evoked responses (ABRs): Prenatal and contemporary correlates. *Child Development, 59*, 571-588.

Murray, A.D. (1988b). Newborn auditory brainstem evoked responses (ABRs): Longitudinal correlates in the first year. *Child Development, 59*, 1542-1554.
Mussen, P.H. (Ed.). (1983) *Handbook of Child Psychology*. New York: Wiley.
Myers, C.S. (1908). Some observations on the development of the colour sense. *British Journal of Psychology, 2*, 353-362.
Naito, T., & Lipsitt, L.P. (1969). Two attempts to condition eyelid responses in human infants. *Journal of Experimental Child Psychology, 8*, 263-270.
Nesselroade, J.R., & Baltes, P.B. (Eds.).(1979). *Longitudinal research in the study of behavior and development*. New York: Academic.
Njiokiktjien, C., & Kurver, P. (1980). Predictive value of neonatal neurological examination for cerebral function in infancy. *Developmental Medicine and Child Neurology, 22*, 736-747.
Norus'sis, M.J. (1986). *SPSS/PC+*. Chicago: SPSS inc.
O'Connor, M.J. (1980). A comparison of preterm and full-term infants on auditory discrimination at four months and on Bayley Scales of infant development at eighteen months. *Child Development, 51*, 81-88.
O'Connor, M.J., Cohen, S., & Parmelee, A.H. (1984). Infant auditory discrimination in preterm and full-term infants as a predictor of 5-year intelligence. *Developmental Psychology, 20(1)*, 159-165.
Olson, G.M., & Sherman, T. (1983). Attention, learning, and memory in infants. In M. Haith & J. Campos (Eds.), *Infancy and the biology of development: Vol. 2* (pp. 1001-1080). In P.H. Mussen (Ed.), *Handbook of child psychology*. New York: Wiley.
Olson, G.M., & Strauss, M.S. (1984). The development of infant memory. In M. Moscovitch (Ed.), *Infant memory* (pp.29-48). New York: Plenum.
Osofsky, J. (Ed.) (1979). *Handbook of infant development*. New York: Wiley.
Oud, J.H., & Sattler, J.M. (1984). Generalized kappa coefficient: A Microsoft BASIC program. *Behavior Research Methods, Instruments, & Computers, 16(5)*, 481.
Papousek, H. (1969). Individual variability in learned responses in human infants. In R.J. Robinson (Ed.), *Brain and Early Behaviour* (pp. 251-266). London: Academic.
Papousek, H., & Papousek, M. (1982). Infant-adult social interactions: Their origins, dimensions and failures. In T.M. Field, A. Huston, H.C. Quay, L. Troll & G.E. Finley (Eds.), *Review of human development* (pp. 148-163). New York: Wiley.
Parmelee, A.H., Kopp, C.B., & Sigman, M. (1976). Selection of developmental assessment techniques for infants at risk. *Merrill-Palmer Quarterly, 22*, 177-199.
Parmelee, A.H., & Michaelis, R. (1971). Neurological examination of the newborn. In J. Hellmuth (Ed.), *Exceptional Infant: Studies in abnormalities: Vol 2* (pp. 3-23). New York: Brunner & Masel.
Pianta, R.C., Sroufe, L.A., & Egeland, B. (1989). Continuity and discontinuity in maternal sensitivity at 6, 24, and 42 months in a high-risk sample. *Child Development, 60*, 481-487.
Porges, S.W. (1979). Developmental design for infancy research. In J. Osofsky (Ed.), *Handbook of infant development* (pp. 742-765). New York: Wiley.

Porges, S.W. (1983). Heart rate patterns in neonates: a potential diagnostic window to the brain. In T. Field & A. Sostek (Eds.), *Infants born at risk. Physiological, perceptual and cognitive processes* (pp. 3-22). New York: Grune & Stratton.
Porges, S.W. (1988). Neonatal vagal tone: Diagnostic and prognostic implications. In P.M. Vietze & H.G. Vaughan (Eds.), *Early identification of infants with developmental disabilities* (pp. 147-159). Philadelphia: Grune & Stratton.
Prechtl, H.F.R. (1982). Assessment methods for the newborn infant, a critical evaluation. In P. Stratton (Ed.), *Psychobiology of the human newborn* (pp. 21-52). New York: Wiley.
Prechtl, H.F.R. (1983). Risk factors and significance of early neurological assessment. In T.B. Brazelton & B.M. Lester (Eds.), *New approaches to developmental screening of infants* (pp. 125-135). New York: Elsevier.
Prechtl, H.F.R., & Beintema, D. (1977). The neurological examination of the full-term born infant. *Clinics in Developmental Medicine*. No 63. London: Heinemann.
Prechtl, H.F.R., & O'Brien, M.J. (1982). Behavioural states of the full-term newborn. The emergence of a concept. In P. Stratton (Ed.), *Psychobiology of the human newborn* (pp. 53-73). New York: Wiley.
Ramey, C.T., & Baker-Ward, C. (1982). Psychosocial retardation and the early experience paradigm. In D.D. Bricker (Ed.), *Intervention with at-risk and handicapped infants* (pp. 269-289). Baltimore: University Park.
Ramey, C.T., Farran, D.C., & Campbell, F.A. (1979). Predicting IQ from mother-infant interactions. *Child Development*, *50*, 804-814.
Reese, A.W., & Lipsitt, L.P. (1970). *Experimental child psychology*. New York: Academic.
Riksen-Walraven, J.M. (1978). Effects of caregiver behavior on habituation rate and self-efficacy in infants. *International Journal of Behavioral Development*, *1*, 105-130.
Rose, S.A. (1981). Lags in the cognitive competence of prematurely born infants. In S.L. Friedman & M. Sigman (Eds.), *Preterm birth and psychological development* (pp. 255-269). New York: Academic.
Rose, S.A. (1983). Differential rates of visual information processing in full-term and preterm infants. *Child Development*, *54*, 1189-1198.
Rose, S.A., Feldman, J.F., McCarton, C.M., & Wolfson, J. (1988). Information processing in seven-month-old infants as a function of risk status. *Child Development*, *59*, 589-603.
Rose, S.A., Feldman, J.F., Wallace, I.F., & McCarton, C.M. (1989). Infant visual attention: Relation to birth status and developmental outcome during the first 5 years. *Developmental Psychology*, *25*, 560-576.
Rosenblith, J.F. (1961a). *Manual for behavioral examination of the neonate as modified by Rosenblith from Graham*. Providence Rhode Island: Brown Duplicating Service.
Rosenblith, J.F. (1961b). The modified Graham behavior test for neonates: test-retest reliability, normative data and hypotheses for future work. *Biology of the Neonate*, *3*, 174-192.
Rosenblith, J.F. (1975). Prognostic value of neonatal behavioral tests. In B.Z. Friedlander, Sterritt, G.M., & Kirk, G.E.(Eds.), *Exceptional Infant: Vol 3. Assessment and Intervention* (pp. 157-172). New York: Brunner Mazel.

Rosenblith, J.F. (1979). The Graham/Rosenblith Behavioral Examination for newborns: prognostic value and procedural issues. In J. Osofsky (Ed.), *Handbook of infant development* (pp. 216-249). New York: Wiley.
Ross, G. (1985). Use of the Bayley Scales to characterize abilities of premature infants. *Child Development*, *56*, 835-842.
Ross, L.E. (1966). Classical conditioning and discrimination learning research with the mentally retarded. In N.R. Ellis (Ed.), *International Review of Research in Mental Retardation: Vol.1* (pp.21-54). New York: Academic.
Ross, L.E., Headrick, M.W., & MacKay, P.B. (1967). Classical eyelid conditioning of young mongoloid children. *American Journal of Mental Deficiency*, *72*, 21-29.
Ross, L.E., & Leavitt, L.A. (1976). Process-research: its use in prevention and intervention with high-risk children. In T.D. Tjossem (Ed.), *Intervention strategies for high-risk infants and young children* (pp. 107-117). Baltimore: University Park.
Ross, L.E., & Ross, S.M. (1973). Classical conditioning and intellectual deficit. In D.K. Routh (Ed.), *The experimental psychology of mental retardation* (pp. 42-77). London: Crosby Lockwood Staples.
Rovee-Collier, C.K., & Fagen, J.W. (1981). The retrieval of memory in early infancy. In L.P. Lipsitt (Ed.), *Advances in infancy research: Vol. 1* (pp. 225-254). Norwood: Ablex.
Rovee-Collier, C.K., & Lipsitt, L.P. (1982). Learning, adaptation and memory in the newborn. In P. Stratton (Ed.), *Psychobiology of the human newborn* (pp. 147-190). Chichester: Wiley.
Rubin, R. A., & Balow, B. (1979). Measures of infant development and socioeconomic status as predictors of later intelligence and school achievement. *Developmental Psychology*, *15*, 225-227.
Ruddy, M.G., & Bornstein, M. H. (1982). Cognitive correlates of infant attention and maternal stimulation over the first year of life. *Child Development*, *53*, 183-188.
Ruff, H.A., McCarton, C., Kurtzberg, D., & Vaughan, H.G.(1984). Preterm infants' manipulative exploration of objects. *Child Development*, *55*, 1166-1173.
Saint-Anne Dargassies, S. (1972). Neurodevelopmental symptoms during the first year of life. *Developmental Medicine and Child Neurology*, *14*, 235-264.
Saint-Anne Dargassies, S. (1977). Long-term neurological follow-up study of 286 truly premature infants. I: Neurological Sequelae. *Developmental Medicine and Child Neurology*, *19*, 462-478.
Saint-Anne Dargassies, S. (1979). Normality and normalization as seen in a long-term neurological follow-up of 286 truly premature infants. *Neuropaediatrie*, *10*, 226-244.
Salapatek, P., & Nelson, C.A. (1985). Event-related potentials and visual development. In G. Gottlieb & N.A. Krasnegor (Eds.), *Measurement of audition and vision in the first year of postnatal life. A methodological review* (pp. 419-453). Norwood: Ablex.
Sameroff, A.J. (1981). Longitudinal studies of preterm infants: A review of chapters 17-20. In S.L. Friedman & M. Sigman (Eds.), *Preterm birth and psychological development* (pp. 387-393). New York: Plenum.

Sameroff, A.J., & Cavanaugh, P.J. (1979). Learning in infancy: A developmental perspective. In J. Osofsky (Ed.), *Handbook of infant development* (pp. 344-392). New York: Wiley.

Sameroff, A.J., & Chandler, M.J. (1975). Reproductive risk and the continuum of caretaking casualty. In F.D. Horowitz (Ed.), *Review of child development research: Vol. 4* (pp. 187-244). Chicago: The University of Chicago.

Sameroff, A.J., & Seifer, R. (1983). Familial risk and child competence. *Child Development, 54*, 1254-1268.

Sattler, J.M. (1988). *Assessment of children* (3rd ed.). San Diego: J.M. Sattler.

Self, P.A., & Horowitz, F.D. (1979). The behavioral assessment of the neonate: an overview. In J. Osofsky (Ed.), *Handbook of infant development* (pp. 126-164). New York: Wiley.

Shotwell, A.M., & Gilliland, A.R. (1943). A preliminary scale for the measurement of the mentality of infants. *Child Development, 14*, 167.

Siddle, D. (Ed.). (1983). *Orienting and habituation*. Chichester: Wiley.

Siegel, L.S. (1982). Reproductive, perinatal, and environmental factors as predictors of the cognitive and language development of preterm and fullterm infants. *Child Development, 53*, 963-973.

Siegel, L.S. (1983). Correction for prematurity and its consequences for the assessment of the very low birth weight infant. *Child Development, 54*, 1176-1188.

Siegel, L.S. (1984). Home environmental influences on cognitive development in preterm and full-term children during the first five years. In A.W. Gottfried (Ed.), *Home environment and early cognitive development* (pp. 197-233). New York: Academic.

Sigman, M. (1983). Individual difference in infant attention: Relations to birth status and intelligence of five years. In T. Field & A. Sostek (Eds.), *Infants born at risk. Physiological, Perceptual and Cognitive Processes* (pp. 271-293). New York: Grune & Stratton.

Sigman, M., & Beckwith, L. (1980). Infant visual attentiveness in relation to caregiver-infant interaction and developmental outcome. *Infant Behavior and Development, 3*, 141-154.

Sigman, M., Cohen, S.E., & Forsythe, A.B. (1981). The relation of early infant measures to later development. In S. Friedman & M. Sigman (Eds.), *Preterm birth and psychological development* (pp. 313-327). New York: Academic.

Smith, A.C., Flick, G.L., Ferriss, G.S., & Sellmann, A.H. (1972). Prediction of developmental outcome at seven years from prenatal, perinatal and postnatal events. *Child Development, 43*, 495-507.

Sokolov, E.N. (1963). *Perception and the conditioned reflex*. New York: MacMillan.

Sophian, C. (1980). Habituation is not enough: Novelty preferences, search, and memory in infancy. *Merrill-Palmer Quarterly, 26*, 239-257.

Spear, N.E., Campbell, B.A. (Eds.). (1979) *Ontogeny of learning and memory*. Hillsdale, NJ: Erlbaum.

St. Clair, K.L. (1978). Neonatal assessment procedures: A historical review. *Child Development, 49*, 280-292.

Stephenson, D, & Siddle, D. (1983). Theories of habituation. In D. Siddle (Ed.): *Orienting and habituation: Perspectives in human research* (pp. 183-236). Chichester: Wiley.

Stratton, P. (1982a). Newborn individuality. In P. Stratton (Ed.), *Psychobiology of the human newborn* (221-261). New York: Wiley.

Stratton, P. (1982b). Significance of the psychobiology of the human newborn. In P. Stratton (Ed.), *Psychobiology of the human newborn* (1-16). New York: Wiley.

Sulzer-Azeroff, B., & Mayer, G.R. (1977). *Applying behavior-analysis procedures with children and youth*. New York: Holt, Rinehart & Winston.

Sweet, A.Y. (1979). Classification of the low-birth infant. In M.H. Klaus & A.A. Fanaroff (Eds.), *Care of the high-risk neonate* (2nd ed.). London: Saunders.

Teller, D.Y. (1979). The forced-choice preferential looking procedure: a psychophysical technique for use with human infants. *Infant Behavior and Development*, *2*, 135-153.

Thomas, H. (1970). Psychological assessment instruments for use with human infants. *Merrill-Palmer Quarterly*, *16*, 179-223.

Tighe, T.J., & Leaton, R.N. (Eds.).(1976). *Habituation: Perspectives from child development, animal behavior, and neurophysiology*. New York: Erlbaum.

Tjossem, T.D. (Ed.). (1976). *Intervention strategies for high-risk infants and young children*. Baltimore: University Park.

Tuber, D.S., Berntson, G.G., Bachman, D.S., & Allen, J.N. (1980). Associative learning in premature hydranencephalic and normal twins. *Science*, *210*, 1035-1037.

Ulvund, S.E. (1984). Predictive validity of assessments of early cognitive competence in light of some current issues in developmental psychology. *Human Development*, *27*, 76-83.

Uzgiris, I.C. (1976). Organization of sensorimotor intelligence. In M. Lewis (Ed.), *Origins of intelligence* (pp. 123-163). London: Wiley.

Uzgiris, I.C., & Hunt, J.M.V. (1975). *Assessment in infancy*. Urbana: University of Illinois.

Valentine, C.W. (1914). The colour perception and colour preferences of an infant during its fourth and eight months. *British Journal of Psychology*, *6*, 363-386.

Van Hasselt, V.B., & Hersen, M. (1987). *Psychological evaluation of the developmentally and physically disabled*. New York: Plenum.

Vietze, P.M. (1988). Do we need another test? In P.M. Vietze & H.G. Vaughan (Ed.), *Early identification of infants with developmental disabilities* (pp. 402-410). Philadelphia: Grune & Stratton.

Vietze, P.M., Abernathy, S.R., Ashe, M.L., & Faulstich, G. (1978). Contingency interaction between mothers and their developmentally delayed infants. In G. P. Sackett (Ed.), *Observing Behavior: Vol.1* (pp. 115-132). Baltimore: University Park.

Vietze, P.M., McCarthy, M., McQuiston, S., MacTurk, R., & Yarrow, L.J. (1983) Attention and exploratory behavior in infants with Down's syndrome. In T. Field & A. Sostek (Eds.), *Infants born at risk* (pp. 251-268). New York: Grune & Stratton.

Vlugt, H. van der (1979). Aspects of normal and abnormal neuropsychological development. In M. S. Gazzaniga (Ed.)., *Handbook of behavioral neurobiology: Vol. II. Neuropsychology* (pp. 99-117). New York: Plenum.
Von Bargen, D.M. (1983). Infant heart rate: a review of research and methodology. *Merrill-Palmer Quarterly, 29,* 115-149.
Wachs, T.D. (1984). Proximal experience and early cognitive-intellectual development: The social environment. In A.W. Gottfried (Ed.), *Home environment and early cognitive development* (pp. 273-328). New York: Academic.
Wachs, T.D., & Gruen, G.E. (1982). *Early experience and human development.* New York: Plenum.
Wachs, T.D., & Smitherman, C.H. (1985) Infant temperament and subject loss in a habituation procedure. *Child Development, 56,* 861-867.
Walberg, A., & Marjoribanks, K. (1976). Family environment and cognitive development: Twelve analytic models. *Review of Educational Research, 46,* 527-551.
Wallace, I.F. (1988). Socioenvironmental issues in longitudinal research of high-risk infants. In P.M. Vietze & H.G. Vaughan (Eds.), *Early identification of infants with developmental disabilities* (pp.356-382). Philadelphia: Grune & Stratton.
Watanabe, K., Miyazaki, S., Hara, K., Kuroyanagi, M., Yamamoto, T., Ito, M., Nakamura, S., & Yamada, H. (1979). Neonatal EEG and computerized tomography. *Neuropaediatrie, 10,* 348-360.
Watson, J.B., & Raynor, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3,* 1-12.
Watson, J.S. (1976). Early learning and intelligence. In M. Lewis (Ed.), *Origins of intelligence* (pp. 199-222). New York: Wiley.
Watson, J.S. (1979). Perception of contingency as a determinant of social responsiveness. In E. B. Thoman (Ed.), *Origins of the infant's social responsiveness* (pp. 33-64) Hillsdale, N.J.: Erlbaum.
Watson, J.S., & Ramey, C.T. (1972). Reactions to response-contingent stimulation in early infancy. *Merrill-Palmer Quarterly, 18,* 219-227.
Werner, J.S., & Siqueland, E.R.. (1978). Visual recognition memory in the preterm infant. *Infant Behavior and Development, 1,* 79-94.
Wilson, W.M. (1987). Age adjustment in psychological assessment of children born prematurely. *Journal of Pediatric Psychology, 12,* 445-450.
Winer, B. J. (1971,2nd ed). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.
Yang, R.K. (1979). Early infant assessment: an overview. In J. Osofsky (Ed.), *Handbook of infant development* (pp. 165-184). New York: John Wiley.
Yarrow, J., & Anderson, B.J. (1979). Procedures for studying parent-infant interaction: A critique. In E. Thoman (Ed.), *Origins of the infant's social responsiveness* (pp.). Hillsdale N.J.: Erlbaum.
Yarrow, L.J., Klein, R.P., Lomonaco, S., & Morgan, G.A. (1975). Cognitive and motivational development in early childhood. In B.Z. Friedlander, B. Sterritt & S. Kirk (Eds.), *Exceptional infant: Vol. 3* (pp. 491-502). New York: Brunner/Mazel.

Zelazo, P.R. (1982). An information processing approach to infant cognitive assessments. In M. Lewis & L.T. Taft (Eds.), *Developmental disabilities. Theory, assessment, and intervention* (pp.229-255). Lancaster: MTP.
Zeskind, P.S. (1983). Production and spectral analysis of neonatal crying and its relation to other biobehavioral systems in the infant at risk. In T. Field & A.M. Sostek (Eds.), *Infants born at risk: physiological and perceptual processes* (pp. 23-43). New York: Grune & Stratton.

Appendix A

OBSERVATION SHEET LEARNING ASSESSMENT

Name: Date: No. Observation: Page:

Column I		Column II		Column III	
Responses	Stimuli	Responses	Stimuli	Responses	Stimuli

Curriculum Vitae

Aloys Goossens werd geboren op 20 maart 1951 te Gemert. Na het eindexamen HBS-B aan het Dr. Knippenberg College (Helmond) studeerde hij natuurkunde en vanaf 1970 psychologie aan de Katholieke Universiteit te Nijmegen. Zijn hoofdrichting was ontwikkelingspsychologie. Vanaf januari 1979 was hij betrokken bij verschillende onderzoeksprojekten van de sectie Zintuiglijk Gehandicapten van het Instituut voor Orthopedagogiek van de Katholieke Universiteit te Nijmegen. Het laatste onderzoeksprojekt (beurs van de Universitaire Onderzoekspool) resulteerde in dit proefschrift.

Sinds 1985 is hij als psycholoog verbonden aan het Instituut voor Doven te Sint Michielsgestel.

STELLINGEN

behorend bij het proefschrift

'Infant Learning Rate within Conditioning Procedures as a Predictor of Subsequent Psychological Development'

1. Klassieke en operante conditioneringsprocedures zijn geschikt om de verwerking van eenvoudige stimuli te beoordelen (dit proefschrift).

2. Op jonge leeftijd zijn klassieke conditioneringsprocedures betrouwbaarder toe te passen dan operante; deze laatsten bezitten echter meer ecologische validiteit (dit proefschrift).

3. Niet-succesvol te conditioneren kinderen vormen een groep die een extra risico loopt met betrekking tot de latere ontwikkeling (dit proefschrift).

4. Evaluatie van ontwikkelingsmogelijkheden door middel van conditionering geeft duidelijker aanknopingspunten voor vroege interventie dan traditionele ontwikkelingstests.

5. Je bent nooit te jong om te leren.

6. In een behandelinstituut is het achterlopen van theoretische kennis op praktische kennis te verkiezen boven het tegenoverstelde.

7. In een zakelijke organisatie wordt winst bereikt door zakelijkheid; in een non-profit organisatie is zakelijkheid winst.

8. Bij het lezen maken oraal-auraal onderwezen dove kinderen gebruik van een articulatorische decodeerstrategie.

9. Restverschijnselen van vroegkinderlijk imitatief gedrag zien we bij volwassenen nog duidelijk terug: wanneer iemand begint met het sproeien van de tuin, sproeit binnen afzienbare tijd de hele buurt.

10. Mijn scepsis ten aanzien van het bestaan van paranormale verschijnselen is in de loop van het voorbereiden van dit proefschrift alleen maar gegroeid: geen van mijn visioenen met betrekking tot het tijdstip van de voltooiing is uitgekomen.

ISBN 90 9003409 9